Constraining Marine Carbon Fluxes

in the Ocean Interior

Jamie D. Wilson

Thesis submitted for the Degree of Doctor of Philosophy

Cardiff University

July 2015

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ………………………………………… (candidate)   Date …………………………

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of …………………………(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed ………………………………………… (candidate)   Date …………………………

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ………………………………………… (candidate)   Date …………………………

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ………………………………………… (candidate)   Date …………………………

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed ………………………………………… (candidate)   Date …………………………

Abstract

The uptake of dissolved CO_2 by phytoplankton in the surface ocean and its delivery to the deep ocean via the remineralisation of sinking particles, the biological pump, is an important control on the exchange of CO_2 between the ocean and the atmosphere. Ocean biogeochemical models suggest that atmospheric CO_2 is sensitive to changes in the depth at which the majority of particles have been remineralised in the ocean interior. However, the key mechanisms involved are not well understood. The function of the biological pump in the past and future is a large uncertainty for the carbon cycle. This thesis uses observations and modelling to further constrain our mechanistic understanding of the biological pump.

Geographically Weighted Regression is applied to an updated sediment trap dataset to explore the spatial variability in statistical relationships between organic matter and $CaCO_3$ that are the basis for the ballast hypothesis. No uniform strong relationship at smaller spatial scales and patterns consistent with surface biogeochemistry suggests ecosystem processes may be important.

In response to the limited sampling of particulate fluxes analysis explored whether annual average fluxes could be estimated from a PO_4 climatology using modelled ocean transport rates in the form of a transport matrix. The Earth System Model GENIE was used to create a synthetic dataset to test this approach, finding significant sources of uncertainty from errors in the observations, the use of modelled transport rates and the assumption that remineralisation is from particles only.

The transport matrix formed a basis for a steady-state phosphorus-only model used to find optimal solutions of spatially varying remineralisation using a 600 member Latin Hypercube ensemble and observed $[PO_4]$. Modelled $[PO_4]$ was predominantly sensitive to global mean remineralisation depths although some spatial variability could be constrained. This has implications for using nutrient distributions to validate mechanistic parameterisations in models.

Author's Note

The chapters in this thesis form the basis for the following publications:

Chapter 3 forms the basis for the following publication:

Wilson, J.D., Barker, S., and Ridgwell, A., (2012) Assessment of the spatial variability in particulate organic matter and mineral sinking fluxes in the ocean interior: Implications for the ballast hypothesis. *Global Biogeochemical Cycles*, 26 (4).

Chapter 4 forms the basis for a paper in review at *Biogeosciences*:

Wilson, J.D., Ridgwell, A., and Barker, S., (2015) Can organic matter flux profiles be diagnosed using remineralisation rates derived from observed tracers and modelled ocean transport rates? *Biogeosciences Discussions*, 12 (6), pp. 4557–4593

Chapter 5 forms the basis for a paper being prepared for submission to *Deep-Sea Research Part I*.

All papers were written by J.D.Wilson with edits contributed by S.Barker and A.Ridgwell.

During the PhD, J.D.Wilson also contributed to the the following publication:

John, E.H., **Wilson, J.D.**, Pearson, P.N., and Ridgwell, A. (2014) Temperature-dependent remineralization and carbon cycling in the warm Eocene oceans. *Palaeogeography, Palaeoclimatology, Palaeocology*, 413 (1), pp. 158 - 166.

J.D.Wilson performed the modelling analysis and wrote sections 3, 4 and 5 with edits from other authors.

Super Mario Bros (1985) Flagpole Fanfare

Original composition for the Nintendo Entertainment System by Koji Kondo
Accurate transcription & optimized fingering for the piano by Joseph Karam

Acknowledgements

I would first like to thank my two supervisors Steve and Andy. They both provided great support in many aspects of the thesis from science to beer (often simultaneously). They put up with my many distractions with non-science projects in MATLAB so as thanks I have coded a MATLAB rendition of the Super Mario Bros Flagpole Fanfare which I'm sure they can't wait to hear! I would also like to thank Andrew Yool and Ian Hall for their insightful and constructive discussion and feedback on the thesis.

I would also like to thank various people who have shared their work, opinions and views that have helped shape this thesis. Thanks to Stephanie Henson, Mathis Hain, Jake Gebbie, Eleanor John, Paul Pearson and Lauren Gregoire. I would also like to particularly thank Sarah Greene, Fanny Monteiro, Sandra Arndt, Ros De'Ath and Sandy Kirtland-Turner who let me share their office in Bristol. Your discussions both about my work and the carbon cycle in general were a real help and encouragement throughout my PhD.

Thanks to everyone in the Wiggly-lines PhD office. Pete originally sent me the link to the transcriptions of the Super Mario Bros soundtrack which have been played on the office Casiotone MT-46 keyboard (49 keys, 8 preset sounds and 8 built-in rhythms!). Sam has always encouraged me to play it whilst it mostly seemed to annoy Scott! Thanks to Milly, Rachel, Elaine and Margit who provided lots of much needed company and tea particularly during finishing. Thanks to Sarah for also being available for tea drinking in Bristol.

Finally, thank you to Cath whose much valued support kept me going throughout.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

Changes in global climate are inherently linked to the concentration of carbon dioxide (CO_2) in the atmosphere by its properties as greenhouse gas. Ice core records of atmospheric CO_2 over the last 800,000 years show that CO_2 concentrations have risen and fallen between 180 and 280 ppm in the quasi-periodic glacial-interglacial cycles in a tight coupling with temperature (*Jouzel et al.*, 2007). In the last century, anthropogenic emissions of CO_2 have led to an increase in atmospheric CO_2 beyond the glacial-interglacial amplitude concurrent with an increase in observed global mean air-surface temperatures (*Brohan et al.*, 2006). Understanding the function of the carbon cycle is key to understanding both the mechanisms behind the glacial-interglacial cycles of CO_2 and changes in atmospheric CO_2 through time as well as understanding the long-term impact and fate of anthropogenic CO_2 emissions.

The amount of carbon in the atmosphere is relatively small ($\sim$600 Pg C; 1 Pg C = 10^{15}g) in comparison to other reservoirs in the carbon cycle and has relatively large exchanges with the biosphere and the surface ocean (Figure 1.1). Consequently atmospheric CO_2 has a small residence time and is very sensitive to changes in the fluxes of carbon in and out of the atmosphere. On timescales of 10-100 years carbon is exchanged between the atmosphere, the biosphere and the surface ocean which combined hold less than 4000 Pg C (Figure 1.1). In comparison, the deep ocean reservoir of carbon is much larger in
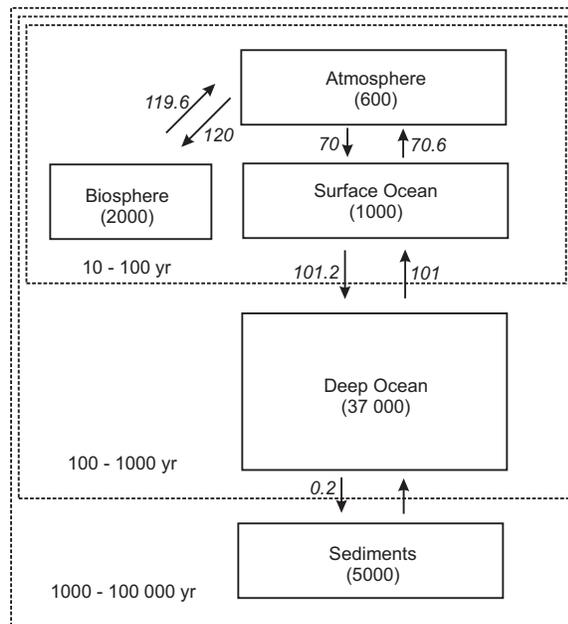
Figure 1.1: Schematic of the pre-indsutrial surface (exogenic) carbon cycle. Approximate reservoir sizes are given in brackets in units Pg C ($1\,\mathrm{Pg} = 10^{15}\mathrm{g}$). Estimated annual fluxes between reservoirs for the 1990s are shown as the italic numbers in units Pg C yr^{-1}. Dashed boxes indicate reservoirs involved in carbon exchange on the timescales indicated. Adapted from *Zeebe and Ridgwell* (2011) with estimated flux values from *Denman et al.* (2007)

size (37,000 Pg C) but it can only interact with the atmosphere via the surface ocean. The timescale of ocean overturning is on the order of 1000 years, setting a limit on the way carbon in the deep ocean can interact with the atmosphere. Because of its size and timescales of interaction with the atmosphere, changes in the marine carbon cycle are thought to be a key factor in the glacial-interglacial cycles (*Sigman and Boyle*, 2000; *Sigman et al.*, 2010) and in the ultimate fate of anthropogenic CO_2 emissions (*Archer et al.*, 2009).

Exchanges between the surface and deep ocean are not only limited to ocean mixing however, as some carbon is actively 'pumped' from the surface ocean to the deep by a process known as the 'biological pump'. This term refers to the net effect of organisms living in the surface ocean taking up carbon and forming either organic matter or calcium carbonate ($CaCO_3$) which then sink into the ocean interior as the particulate remains of the organisms. The organic matter component of biological pump is esti-

mated to export 5-10 Pg C per year from the surface ocean to the deep ocean (*Henson et al.*, 2011) and helps maintain a vertical gradient of carbon in the ocean, in part facilitating the uptake of atmospheric CO_2 in the surface ocean. However, the precise mechanisms controlling the biological pump, and their potential role in past and future changes in climate, are uncertain. This introductory chapter of the thesis starts by introducing the biological pump and its interaction with the marine carbon cycle (Section 1.1) in the context of the biological pump as a geochemical process, *i.e.,* considering its net effect on ocean chemistry and how changes in it can affect atmospheric CO_2, such as *Sigman et al.* (2010). Section 1.2 then focuses on observations of the biological pump in the modern ocean and current hypotheses about the key mechanisms driving the biological pump. These are related back to the discussion in Section 1.1 and their implications for the role of the biological pump in changing atmospheric CO_2 through time. Finally, the aims and outline of the thesis are presented.

1.1   The Biological Pump and the Carbon Cycle

1.1.1   Factors Controlling $p\text{CO}_2$ in the Surface Ocean

The exchange of CO_2 between the surface ocean and the atmosphere is dependent on the difference between the partial pressure of CO_2 ($p\text{CO}_2$) in the atmosphere and surface ocean. Assuming no change in fluxes from other reservoirs in the carbon cycle, changes in surface ocean $p\text{CO}_2$ will drive changes in atmospheric CO_2. $p\text{CO}_2$ in the surface ocean is determined by the concentration of dissolved CO_2 and the solubility of the water determined by the coefficient K_0:

$$p\text{CO}_2 = \frac{\text{H}_2\text{CO}_3^*}{\text{K}_0} \tag{1.1}$$

Dissolved CO_2 is referred to as the hypothetical acid H_2CO_3^* because when gaseous CO_2 dissolves into seawater it hydrates into aqueous CO_2 which further reacts to form
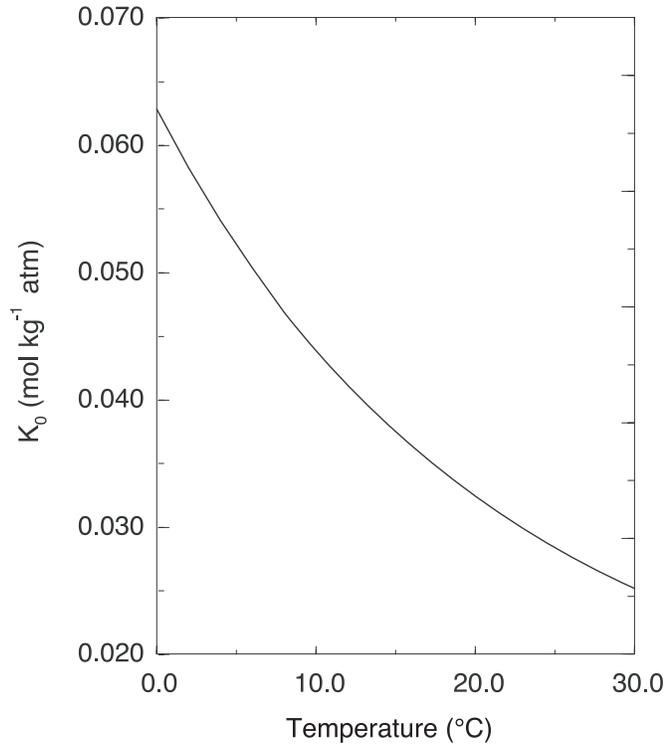
Figure 1.2: Sensitivity of K_0 to temperature. For a constant $[H_2CO_3^*]$, pCO_2 will vary predominantly with temperature as per equation 1.1. Figure adapted from *Sarmiento and Gruber* (2006)

.

carbonic acid which are not chemically separable from eachother (*Zeebe and Wolf-Gladrow*, 2001). K_0 is particularly sensitive to temperature (Figure as well as salinity but the controls on pCO_2 are more complicated due to the inorganic chemistry of carbon in seawater than suggested by equation 1.1.

$H_2CO_3^*$ dissociates into bicarbonate (HCO_3^-) and carbonate (CO_3^{2-}) forming a thermodynamic equilibrium:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3^* \rightleftharpoons HCO_3^- + H^+ \rightleftharpoons CO_3^{2-} + 2H^+ \tag{1.2}$$
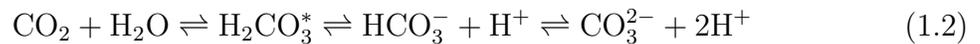
The inorganic carbon system is commonly described by two parameters: dissolved inorganic carbon (DIC), the sum of the carbon species in equation 1.2, and total alkalinity (ALK), a measure of the total bases (proton acceptors) over acids (proton donators)

(*Zeebe and Wolf-Gladrow*, 2001):

$$DIC = [H_2CO_3^*] + [HCO_3^-] + [CO_3^{2-}] \tag{1.3}$$

$$ALK = [HCO_3^-] + 2[CO_3^{2-}] + [OH^-] + [B(OH)_4^-] - [H^+] + \text{minor bases} \tag{1.4}$$

where $[OH^-]$ is the concentration of the hydroxide ion and $[B(OH)_4^-]$ is the concentration of the borate ion. Alternatively, alkalinity can be viewed as the charge balance between all strong acids and bases including nitrate ($[NO_3^-]$). DIC and ALK contain common terms because although an increasing concentration of dissolved CO_2 will increase DIC, the final amount of dissolved CO_2 is also dependent on current state of the inorganic carbon system. Alkalinity indicates how the state of inorganic carbon system with respect to its equilibrium. Additional parameters of the inorganic carbon system such as pH, total dissolved boron and the concentration of each carbon species can also be used. Global mean surface DIC in the modern ocean is dominated by 86.6% $[HCO_3^-]$ and 10.9% $[CO_3^{2-}]$ with $[H_2CO_3^*]$ forming only 0.5%. Likewise, ALK is dominated by the anions in DIC, $[HCO_3^-]$ at 76.8% and $[CO_3^{2-}]$ at 18.8% (*Zeebe and Wolf-Gladrow*, 2001; *Sarmiento and Gruber*, 2006).

As $H_2CO_3^*$ contributes a small fraction of the DIC pool, changes in surface ocean pCO$_2$ are ultimately driven by changes in HCO_3^- and CO_3^{2-}. Therefore, pCO$_2$ in the surface ocean is a function of temperature, salinity as well as DIC and ALK (*Sarmiento and Gruber*, 2006). In the modern ocean, average concentrations of DIC and ALK are 15% and 5% higher respectively in the deep ocean than at the surface (*Sarmiento and Gruber*, 2006). Approximately 90% of this vertical gradient is estimated to be maintained by the action of biology via the biological pump (*Gruber and Sarmiento*, 2002).

1.1.2   The Biological Pump

The biological pump is formed of two individual components, the 'soft-tissue pump' and the 'carbonate pump', that interact with $p\text{CO}_2$ via DIC and ALK in different ways (Figure 1.3).

1.1.2.1   Soft-Tissue Pump

In the upper sunlit zone of the ocean (0-100m), a range of photosynthetic organisms, such as diatoms, cocclithophores, and other picophytoplankton, use sunlight to split water molcules and fix carbon from dissolved CO_2 during photosynthesis to form the carbohydrates in organic matter. Other nutrients such as phosphate (PO_4^{3-}) and nitrate (NO_3^-) are also taken up from solution during photosynthesis (*Sarmiento and Gruber*, 2006):

$$106\text{CO}_2 + 16\text{HNO}_3^- + \text{H}_3\text{PO}_4^{2-} + 78\text{H}_2\text{O} + 18\text{H}^+ \rightleftharpoons \text{C}_{106}\text{H}_{175}\text{O}_{42}\text{N}_{16}\text{P} + 150\text{O}_2 \quad (1.5)$$
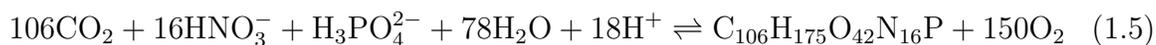
Upon the death of these organisms, a large proportion of the organic matter that was created in the surface ocean is broken down during heterotrophic respiration by zooplankton and bacteria *in-situ, i.e.,* in the surface, releasing CO_2 and nutrients back into solution. However a fraction of the organic matter (0-25% of which is regionally variable; *Henson et al.*, 2011) survives respiration in the surface and is exported into the ocean interior as either sinking particles (particulate organic matter/carbon: POM/POC) or smaller suspended particles (dissolved organic matter/carbon: DOM/-DOC; this is defined operationally as organic matter passing through a 0.45 μm filter *Williams and Follows*, 2011) (Figure 1.3). Estimates suggest that between 5-12 Gt C year^{-1} is exported from the surface to the ocean interior (*Laws et al.*, 2000; *Dunne et al.*, 2007; *Lutz et al.*, 2007; *Henson et al.*, 2011), with DOC forming approximately 20% of
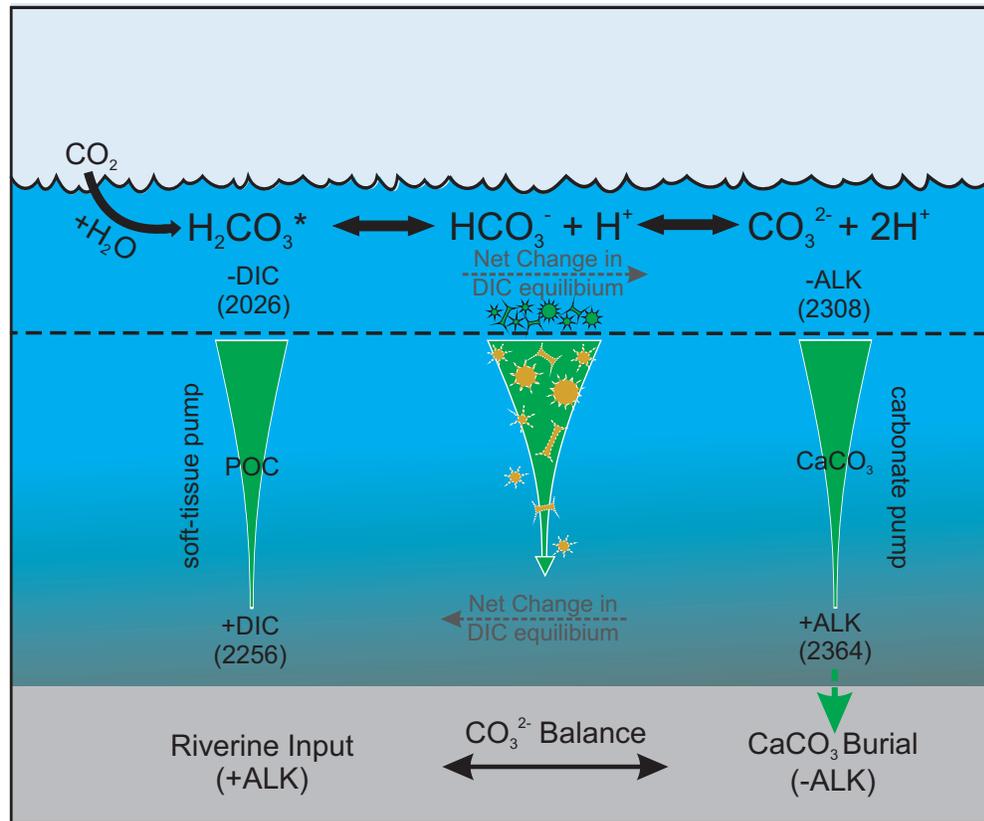
Figure 1.3: Schematic of the biological pump showing the removal of carbon and alkalinity from the surface ocean (above dashed line) to the ocean interior (below dashed line). The dominant effect on ocean chemistry in terms of DIC and ALK for are shown above and below each pump. The average concentrations of surface and deep DIC and ALK are indicated in brackets in units μM. The net effect of the combined pumps on the thermodynamic equilibrium in the modern ocean is shown in grey. The controls on the balance of $[CO_3^{2-}]$ are shown at the bottom of the schematic. Adapted from *Hain and Sigman* (2014) and *Ridgwell and Arndt* (2014)

.

this export (*Hansell et al.*, 2009). Ultimately, the amount of organic carbon exported into the ocean interior ('export production') as opposed to the total organic carbon fixed at the surface ('net primary production') is the important factor for altering the exchange of atmospheric CO_2 (*Sarmiento and Gruber*, 2006).

POM sinking below the surface is further consumed by bacteria and zooplankton as it sinks through the ocean interior, returning CO_2 and nutrients back into solution at depth. This net process of returning CO_2 back into solution is termed 'remineralisation'. The majority of organic matter (>90%; *Francois et al.*, 2002) is remineralised by

<2000m water depth leaving a small fraction of the POM originally exported from the surface reaching the deep ocean (*Martin et al.*, 1987).

The soft-tissue pump has a direct effect on surface ocean pCO$_2$ by maintaining a vertical gradient in DIC, due to the removal of CO$_2$ at the surface and delivery to depth, in the face of ocean circulation which is acting to constantly erode the gradient via mixing. The removal of NO$_3^-$ ions associated with the export of POM increases ALK at the surface and decreases it at depth with remineralisation, although this effect is much smaller due to the amount of NO$_3^-$ taken up during photosynthesis relative to carbon *e.g.,* a global average molar ratio across phytoplankton of 16:117 (*Anderson and Sarmiento*, 1994). Other forms of nitrogen, such as ammonium, can also taken up during the formation of organic matter but are not considered in this thesis.

1.1.2.2   Carbonate Pump

As organic matter is being formed, some photosynthetic organisms, such as coccolithophores and zooplankton such as foraminifera, also precipitate hard parts formed of the mineral calcium carbonate (CaCO$_3$):

$$Ca^{2+} + 2HCO_3^- \rightleftharpoons CaCO_3 + CO_2 + H_2O \tag{1.6}$$

The export of CaCO$_3$ to the ocean interior at $\sim$100m depth is estimated to be 6-8% of the organic carbon export (*Sarmiento et al.*, 2002; *Jin et al.*, 2006). The formation of CaCO$_3$ at the surface has the net effect of decreasing ALK by removing 2 moles of HCO$_3^-$ but also increasing DIC by the addition of 1 mole of CO$_2$ (partly balancing the removal of 2 moles of HCO$_3^-$ from DIC). As with the soft-tissue pump, some of these particles sink into the ocean interior and partly dissolve leading to the reverse reaction (Equation 1.6). The carbonate pump sets up the vertical gradient of ALK in the ocean which also has a direct effect on pCO$_2$ but in the opposite direction to the soft-tissue

pump.

Unlike the remineralisation of POM, the dissolution of sinking $CaCO_3$ particles is sensitive to the saturation state of seawater with respect to $CaCO_3$. The solubility product for the two main forms of $CaCO_3$ is defined as:

$$K_{sp}^{CaCO_3} = [CO_3^{2-}]_{sat}[Ca^{2+}]_{sat} \tag{1.7}$$

$K_{sp}^{CaCO_3}$ is strongly dependent on pressure leading to undersaturation at depth, although this varies between basins due to changes in $[CO_3^{2-}]$ with the soft-tissue pump (*Sarmiento and Gruber*, 2006). Ca^{2+} forms a major constituent of modern seawater and is often assumed a constant (*Sarmiento and Gruber*, 2006). A larger fraction of the exported $CaCO_3$, compared to exported POC, reaches the seafloor and is involved in interactions with the sediments.

1.1.2.3   Interactions with Sediments

Sinking particulates of organic matter and $CaCO_3$ that reach the seafloor are incorporated into the ocean sediments (Figure 1.3). Interactions between the two components of the biological pump and the sediments are also an important control on pCO_2 via whole ocean changes in DIC and ALK. The sediments consist of solid particles made up of clays, $CaCO_3$, silica, and clays, that derive from the sinking particles at the sediment surface and form a matrix filled with pore water (*Sarmiento and Gruber*, 2006). DIC and ALK in the sediment porewater can change as a function of the remineralisation of organic matter particles and the dissolution of minerals. These processes lead to large concentration gradients between the pore water and overlying seawater leading to the diffusion between the two, as well as within different layers of the sediments. As particulates rain onto the sediment surface, consolidation of the sediment below and movements of the ocean crust contribute to a downward motion of the sediments, lead-

ing to some particles becoming ultimately buried. The sediment processes that have the greatest impact on the dissolved constituents of the ocean occur in the top $\sim$10cm of the sediment (*Sarmiento and Gruber*, 2006) and are considered further.

The precipitation of $CaCO_3$ in the surface ocean and burial in the sediments via the carbonate pump represents a net loss of $[CO_3^{2-}]$ and therefore ALK from the ocean over time. The weathering of rocks on land also delivers $[CO_3^{2-}]$ and ALK to the ocean via rivers representing a input of ALK in the ocean. By acting as a control on the loss of ALK, changes in the carbonate pump can have a direct effect on the balance of ALK in the ocean (*Ridgwell and Zeebe*, 2005). Equally, changes in the soft-tissue pump can alter the burial of $CaCO_3$ in the sediments by altering the concentration of dissolved CO_2 in the deep ocean driving a change in the saturation state of the deep ocean affecting the dissoution of $CaCO_3$ in the surface sediments, *e.g.,* equation 1.7 (*Archer*, 1991; *Arndt et al.*, 2013). The majority of organic matter reaching the sediment is remineralised (*Sarmiento and Gruber*, 2006; *Arndt et al.*, 2013) increasing the concentration of CO_2 in the pore water and increasing the solubility of $CaCO_3$, providing an additional mechanism for the soft-tissue pump to alter ocean ALK. In summary, any net change in the ratio of POC to $CaCO_3$ in the deep ocean and sediment surface will drive changes in the sediment composition and the ALK of the ocean.

1.1.3   How can changes in the Biological Pump drive changes in Atmospheric CO_2?

The exchange of CO_2 between the atmosphere and ocean, and therefore the concentration of CO_2 in the atmosphere, is sensitive to changes in the biological pump due to its effects on DIC and ALK. This section categorises and quantifies the potential ways in which the two components of the biological pump could change and affect atmospheric CO_2. These are broadly categorised as changes occurring in the surface ocean and changes occurring in the ocean interior in order to highlight that changes in the ocean

interior are potentially significant for atmospheric CO_2 but have not yet been considered in as much detail. The following discussion draws on concepts and hypotheses about the role of the biological pump in glacial-interglacial cycles, which has been a focus of research when considering the biological pump as a geochemical process *e.g., Sigman and Boyle* (2000); *Sigman et al.* (2010); *Hain and Sigman* (2014).

The removal of DIC and ALK from the surface ocean and return at depth through the action of the biological pump maintains vertical gradients in DIC and ALK in the face of ocean circulation. The circulation is constantly acting to erode these gradients and reverse the effect the biological pump by returning DIC and/or ALK back to the surface ocean (*Hain and Sigman*, 2014). Changes in atmospheric CO_2 can occur by increasing/decreasing the amount of DIC and ALK in the ocean interior. For example, a relative increase of DIC in the ocean interior will lead to a drawdown of CO_2 whereas an increase in ALK will lead to an increase in atmospheric CO_2. This concept can be used to define an "efficiency" of the biological pump. As the soft-tissue pump is limited by nutrients at the surface, *e.g.,* the greater proportion of nutrients maintained in the deep ocean, the greater the efficiency of the biological pump. In the modern ocean, a proportion of the nutrients bought to the surface are not utilised by biology before being subducted back into the ocean interior, representing a "missed opportunity" for the biological pump (*Sigman et al.*, 2010). Estimates of these nutrients ('preformed nutrients') based on oxygen utilisation in the modern ocean suggest that between 35% to 50% of nutrients upwelled to the surface ocean are actually utilised by biology (*Ito and Follows*, 2005; *DeVries et al.*, 2012a; *Duteil et al.*, 2013). Modelling experiments have shown that atmospheric CO_2 is very sensitive to changes in the global concentration of preformed nutrients in the ocean (*Ito and Follows*, 2005; *Marinov et al.*, 2008). The efficiency of the biological pump can therefore be changed by changing the rate of nutrients delivered to the surface or by changing the rate at which nutrients are taken up by biology. In the following sections, changes in the biological pump efficiency driven by biological processes are described. *Sigman and Boyle* (2000), *Sigman et al.* (2010)
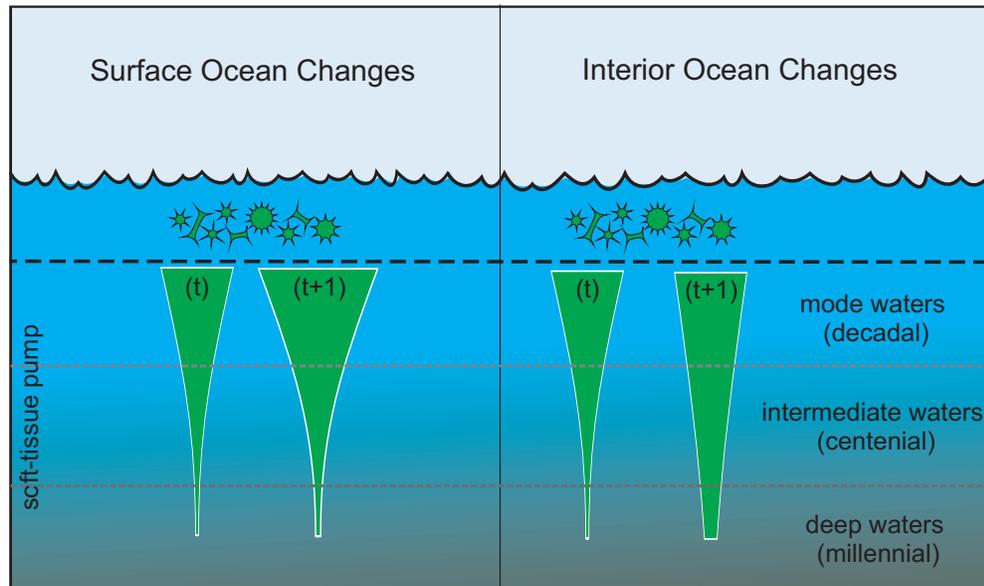
Figure 1.4: Schematic depicting two ways in which biological driven changes in the biological pump could decrease atmospheric CO_2 from abitrary time t to t+1. Surface ocean changes change the total amount of carbon exported to the deep ocean whereas interior ocean changes affect the timescale at which carbon is returned to the surface by circulation. Text indicates the depth of the ocean where POC is remineralised and the approximate timescale it remineralised carbon is returned to the surface by circulation. Adapted from *Boyd and Trull* (2007).

and references therein consider changes in the biological pump alternatively driven by changes in circulation in detail which are not discussed further in this thesis.

1.1.3.1   *Changes in the Surface Ocean*

Changes in the amount of carbon exported from the surface ocean to the ocean interior have been a major focus for previous research and could occur a number of different ways: changes in the magnitude of carbon exported and changes between the relative contribution of the two component pumps to export. In large parts of the low and mid latitudes of the modern ocean, phytoplankton utilise the majority of nutrients available. The amount of carbon pumped into the ocean interior is therefore ultimately limited by the availability of nutrients. *Broecker* (1982) proposed that increases in the nutrient inventory (PO_4 and NO_3^-) of the ocean could increase the total amount of productivity and carbon exported to the deep ocean leading to an uptake of CO_2. A 30% increase in

the nutrient inventory is predicted to lead to a 17 - 46 ppm drawdown of CO_2 dependent on whether there is an equal response in the carbonate pump (*Sigman and Boyle*, 2000).

In the modern ocean, the net nutrient utilisation by phytoplankton at the surface is not complete everywhere. In the Southern Ocean, deep nutrient-rich water is upwelled to the surface and flows equatorward before being subducted into mid-depths still containing a proportion of the original upwelled nutrients (*Ito and Follows*, 2005; *Sarmiento and Gruber*, 2006). Despite high nutrient concentrations, productivity is relatively low leading to this region, as well as others, being termed High Nutrient Low Chlorophyll (HNLC) regions. Productivity in HNLC regions is thought to be low due to either zooplankton grazing suppressing phytoplankton communities, the limitation of a micronutrient such as iron, or a combination of both (*Sarmiento and Gruber*, 2006). An increased flux of iron to the Southern Ocean associated with increased dust deposition during the Last Glacial Maximum has been hypothesised to increase producitivity in this region and increase the biological pump efficiency (*Martin*, 1990; *Ziegler et al.*, 2013). Reductions in the volume of water ventilated to the Southern Ocean has also been a significant focus, see *Sigman et al.* (2010) and *Hain and Sigman* (2014).

Changes in the biological pump in the surface ocean can also reflect relative changes between the soft-tissue and carbonate pumps. *Archer and Maier-Reimer* (1994) demonstrated that a 40% decrease in the 'rain-ratio' between $CaCO_3$ and POC (from 0.25 to 0.15) exported from the surface could be responsible a 70-90ppm decrease in atmospheric CO_2 via interactions with deep-sea sediments. The drawdown of atmospheric CO_2 was largely a function of including the dissolution of $CaCO_3$ within sediments due to *in-situ* remineralisation of organic carbon in sediment pore-waters (*Archer and Maier-Reimer*, 1994). The increased dissolution of sedimentary $CaCO_3$ leads to an increase in $[CO_3^{2-}]$ in the ocean, driving a reduction in CO_2 in the atmosphere.

1.1.3.2   Changes in the Ocean Interior

More recently, there has been an increased focus on the processes of the biological pump in the ocean interior that have previously been assumed to be relatively static. Changes in the rate of remineralisation of sinking POC or the rate at which it sinks can potentially change the depths at which CO_2 is returned to solution. In general, the deeper organic matter is remineralised in the ocean interior, the longer it takes for the dissolved CO_2 to reach the surface ocean (*Boyd and Trull*, 2007) (Figure 1.4). By reducing the rate at which nutrients return to the surface, the surface ocean has a relatively lower concentration of DIC leading to drawdown of atmospheric CO_2. *Kwon et al.* (2009) measured the sensitivity of atmospheric CO_2 to changes in the average depth of POC remineralisation in an ocean biogeochemical model. An increase of 23m from 204m to 228m in the *e*-folding depth (the depth at which 63% of the exponentially decreasing particulate organic matter flux has been remineralised) was associated with a 10-27ppm drop in atmospheric CO_2, 85% of which was due to the soft-tissue pump (*Kwon et al.*, 2009).

Although the rain-ratio mechanism as proposed by *Archer and Maier-Reimer* (1994) described above, focuses on changes in export, the ultimate reason it affects atmospheric CO_2 is due to relative changes between the organic matter and carbonate fluxes at the sediment surface (*Ridgwell*, 2003). This is because the mechanism relies on the dissolution of $CaCO_3$ driven by organic matter remineralisation in the sediment porewater (*Archer and Maier-Reimer*, 1994). Any change in the delivery of POC or $CaCO_3$ can therefore have an impact on atmospheric CO_2 via interactions with deep-sea sediments. Changes in the remineralisation of POC as described above, could also lead to changes in the organic matter reaching the sediments relative to $CaCO_3$. A model sensitivity study found that the atmospheric CO_2 drawdown associated with the same perturbation used by *Kwon et al.* (2009), was up to 4 times higher when including sediment feedbacks (8.1 ppm *cf.* 33.4 ppm) (*Roth et al.*, 2014). The initial drawdown of CO_2 from

the redistribution of DIC and ALK (see *Kwon et al.*, 2009) occurred over $\sim$2000 years whilst taking $\sim$50 000 years to reach a new equilibrium (*i.e.,* the inventory of $[CO_3^{2-}]$ reached a new steady state value) when including sediment feedbacks. In comparison, atmospheric CO_2 was much less sensitive to changes in the $CaCO_3$ remineralisation depth, lowering by 1.8 ppm for the closed system and 12.2 ppm when perturbed by the same amount (*Roth et al.*, 2014).

1.2   Mechanistic Understanding of the Biological Pump in the Ocean Interior

As discussed in the previous section, the concentrations of CO_2 in the atmosphere is potentially sensitive to relatively small changes in the efficiency of the soft-tissue component of the biological pump. Despite this, the efficiency of the biological pump has commonly been considered static in time and space (*Boyd and Trull*, 2007). Observations of sinking particles in the modern ocean have led to speculation about potential mechanisms controlling the remineralisation of particles in the ocean interior that could have implications for understanding how the biological pump has changed in the past and will interact with current changes in the ocean due to climate change. In this section modern observations and existing mechanistic hypotheses on the controls of sinking particulate organic matter are described with a focus on the implications for our understanding of the carbon cycle described in the previous section.

1.2.1   Observations of the Biological Pump

1.2.1.1   Sediment Traps

Global observations of sinking particulate material in the ocean interior predominantly come from sediment traps: devices that trap sinking particles using a funnel that covers a set area, *e.g.,* $0.25m^{-2}$ (*WHOI*, 2007) and which preserve the samples to protect them

against degradation by biological processes (Figure 1.5). Traps, such as the commonly used design illustrated in Figure 1.5a, can be moored in the deep ocean, often for 12 month periods, with the ability to take multiple samples throughout the period. *Honjo et al.* (2008) collated a global dataset of 134 annually resolved sediment trap observations from the deep ocean ($\sim$2km) over the last three decades using this design.

Sediment traps can be subject to a number of biases. Sinking particles that are entering the trap, as well as those already within the trap, can be entrained by eddies created by water flowing across the surface of the trap causing the trap to under-sample the flux (*Buesseler et al.*, 2007). This bias is particularly significant for shallower surface-tethered sediment traps (Figure 1.5b) where flow rates are generally higher. Neutrally-buoyant sediment traps have been recently developed to address this issue (Figure 1.5b; *Buesseler et al.*, 2007). Zooplankton can also enter the trap, contributing additional organic material or changing elemental ratios through their excretions and consumption of the organic matter. This bias is concentrated in the shallower ocean where zooplankton are found in greater numbers (*Buesseler et al.*, 2007). Finally, once sampled, chemical elements in the particles can become soluble, entering the supernatant in the sample bottle, *e.g., Brzezinski and Nelson* (1995). Analysis of the supernatant can be used to correct for this. A globally distributed dataset of deep sediment trap data produced with consistent methodology is available as part of the Joint Global Ocean Flux Study (JGOFS: *Honjo et al.*, 2008). This dataset comprises of samples from 134 sediment trap sites which although are globally distributed, are not distributed equally with large regions of the ocean such as southern subtropical and temperate regions of all basins missing any samples. Of the 134 samples, 47 are in the Pacific, 50 in the Atlantic, 18 in the Southern Ocean, and 11 in the Indian Ocean (see Figure 3.1). A new approach to generating high resolution estimates of POC flux curves is explored in Chapter 4.
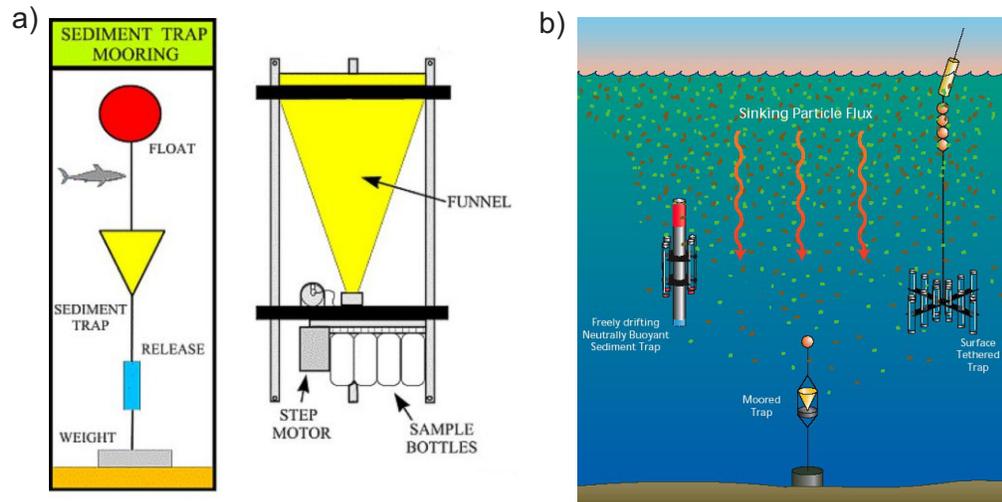
Figure 1.5: Schematics of sediment traps and their deployments. (a) the design of a moored time-series sediment trap (Mark 78G-21) by McLane Research Laboratories similar to the traps used in the Joint Global Ocean Flux Study (*Honjo et al.*, 2008). Figure adapted from *USGS* (2004). (b) Schematic of the range of sediment trap designs used. Figure adapted from *WHOI* (2007).

1.2.1.2   *Other Observations*

Observations of POC fluxes are also available from other methods. The concentration of POC in the water column can be measured by filtering thousands of litres of seawater in-situ using a Multiple Unit Large Volume Filtration System (MULVFS) (*Lam et al.*, 2011). POC fluxes are then estimated by considering sinking speeds and particles characteristics (*Lam et al.*, 2011). The MULVFS has the advantage of being able to take samples with more vertical resolution in the mesopelagic ocean (0-1000m) without the biases of shallow sediment traps. *Lam et al.* (2011) provide an assessment of the globally distributed data collected to date.

1.2.1.3   *Global-Scale Estimates of POC Flux Profiles*

Vertical profiles of POC fluxes from sediment trap observations show that particle fluxes reduce rapidly with depth as POC is remineralised until 1000-1500m, below which fluxes are sinking without significant modification (Figure 1.6a: *Martin et al.*, 1987). On average the flux in the deep ocean (>2000m) is ~5% of the original flux

being exported (*Armstrong et al.*, 2002). Vertical POC fluxes are commonly described empirically by fitting a power-law function to sediment trap data:

$$\frac{dF_z}{dz} = F_{z0} \cdot \left(\frac{z}{z0}\right)^b \tag{1.8}$$

where F_z is the flux at depth z, $z0$ is the depth of the bottom of the surface layer and b is a dimensionless parameter that controls the attenuation of the flux profile with depth. A power-law curve was first found to be a best fit to 48 open ocean sediment trap observations in the North-East Pacific by *Martin et al.* (1987) who found a value of b=-0.858±0.1, a value which has been used extensively in subsequent data and modelling studies. A power-law curve with this value of b is known as the 'Martin Curve'. Various other functions and algorithms have also been used to describe flux profiles including other power-law functions (*Suess*, 1980; *Pace et al.*, 1987) and exponential functions (*Armstrong et al.*, 2002; *Andersson et al.*, 2004; *Lutz et al.*, 2002, 2007) (see Figure 5.1 for example). It is worth noting that the choice of function reflects certain assumptions about the relationship between sinking speeds and remineralisation rates. An exponential function represents a fixed remineralisation rate and sinking speed, whereas a power-law curve represents an increasing sinking speed or equally a decreasing remineralisation rate with depth (*Kriest and Oschlies*, 2008; *Lam et al.*, 2011).

Compilations of sediment trap data have calculated that the global average of $b =$ -0.639 (*Henson et al.*, 2012), -0.70±0.08 (*Primeau*, 2006) and -0.82±0.16 (*Berelson*, 2001), which are generally larger (deeper remineralisation) than that found by (*Martin et al.*, 1987). The flux curve has been widely used to model POC fluxes in ocean biogeochemical models. Several studies have systematically varied the exponent b and compared the resulting tracer fields to observations to infer that the optimal global value of $b = $ -0.9 (*Yamanaka and Tajika*, 1996), -1.0 (*Kwon and Primeau*, 2006), ∼-0.858 (*Kriest et al.*, 2012). These inferred values are generally consistent with the value
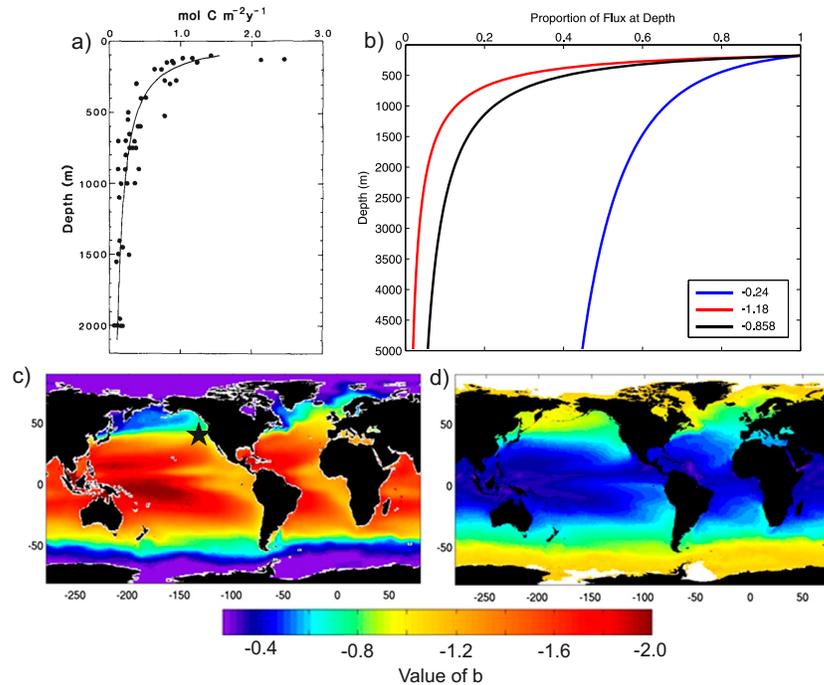
Figure 1.6: Comparison of POC flux profiles derived from sediment trap data. (a) The original 'Martin Curve' with a value of b=-0.858. From *Martin et al.* (1987). (b) A comparison of the range of flux curves versus the Martin Curve with values for b from *Henson et al.* (2012) using an export depth of 100m. (c) The spatial variability of b from *Henson et al.* (2012). The black star indicates the approximate location of the data used to fit the Martin Curve in panel a. (d) The spatial variability of b from shallower sediment traps found by *Marsay et al.* (2015). White areas indicate where chlorophyll data from satellite measurements is unavailable.

found by (*Martin et al.*, 1987).

Subsequent studies of globally distributed sediment trap observations have suggested that the value of b may vary regionally. Values of b <-0.858 indicate that more remineralisation is occurring higher in the water column, and b >-0.858 indicates greater remineralisation occurring at depth. A study of regional sediment trap profiles found values of b varying regionally between -1.28 and -0.59 (*Berelson*, 2001) although *Primeau* (2006) suggests statistical errors may contribute to this range. More recently, *Henson et al.* (2012) performed a systematic comparison of primary production, export production and flux algorithms applied to a satellite based climatology of chlorophyll. Estimates of POC fluxes were compared with sediment trap observations to find the best set of algorithms which were then used to produce a high resolution field of b values

ranging from -1.18 to -0.24 (Figure 1.6c: *Henson et al.*, 2012). The spatial variability in b produced by this approach is predominantly latitudinal with shallower remineralisation at high latitudes ($b <$-0.858) and deeper remineralisation in low latitude regions ($b >$-0.858) (Figure 1.6b and Figure 1.6c). Alternative measures, such as the 'transfer efficiency', defined as the fraction of flux typically remaining at depth ($\sim$2000m), also confirm this large-scale spatial pattern in POC fluxes (*Francois et al.*, 2002; *Lutz et al.*, 2002; *Honjo et al.*, 2008).

These studies are based on deep sediment traps, *e.g.,* $>$1000-1500m, due to the sampling biases of shallower traps (*Buesseler et al.*, 2007). A recent compilation of neutrally-buoyant sediment traps from shallower depths ($<$1000m) has suggested that the spatial pattern of POC flux profiles is the inverse of those previously found (Figure 1.6d; *Marsay et al.*, 2015). The authors suggest the result may be a function of different components of the POC flux that are decoupled from eachother. However, this study does not contain data from regions such as the Southern Ocean (Figure 1.2.1) although this region is also relatively undersampled in the previous studies (Figure 1.6). Modelling studies have not fully applied the approach of optimising regionally variable values of b in ocean biogeochemical models (although see *Yao and Schlitzer*, 2013) which is explored further in Chapter 5.

1.2.2   Hypothesised Mechanisms

Ultimately, the regional variability observed in b, or more generally in the depth of remineralisation as described empirically by a mathematical function, represents a range of potential processes such as the particle dynamics, the bacterial breakdown of organic matter, zooplankton grazing, and the effect of minerals that could contribute to the vertical attenuation of POC fluxes. The following section describes current hypotheses about the mechanisms and processes leading to the observed spatial variability and their implications for the carbon cycle.

1.2.2.1   Ballasting

The synthesis of globally distributed sediment trap data from the deep ocean (>1000-1500m; where undersampling biases due to currents are considered minimal) provided the basis for the 'ballast' hypothesis Despite large variability in the absolute magnitude of POC fluxes exported from the surface, the dry weight of POC at depth, when normalised to the total particulate flux, converges globally to approximately 5% with significant regional variability. The majority of the total particulate flux ($\sim$90%) at depth is however made up of minerals such as $CaCO_3$ and opal rather than POC. This was used as a basis for a mechanistic model that predicts the flux of POC at depth (F_{POC}) as a function of POC:mineral ratios by *Armstrong et al.* (2002). The model is formulated as an exponential model (*Armstrong et al.*, 2002):

$$F_{POC(z)} = F_{POC(\infty)} + (F_{POC(z0)} - F_{POC(\infty)}) \cdot \exp^{-(\frac{z-z0}{l_{POC}})} \qquad (1.9)$$

where a fraction of the POC flux ($F_{POC(z0)} - F_{POC(\infty)}$) that is un-associated with minerals remineralises according to a globally constant length scale (l_{POC}) similar to the Martin Curve. The deep asymptotic flux ($F_{POC(\infty)}$) is determined by the POC:mineral ratio. *Armstrong et al.* (2002) suggested that the quantitative associated between POC and minerals at depth may be due to the physical protection of some POC from remineralisation by being incorporated into the mineral structure as well as POC acting as a glue to bind particles together (*Armstrong et al.*, 2002).

Expanding on the variability in POC:mineral ratios, *Klaas and Archer* (2002) applied multiple linear regression to annual average flux data from a globally sampled set of 78 deep sediment traps to investigate the interactions between POC and the different types of minerals $CaCO_3$, opal and 'lithogenic' material: inorganically derived minerals from sources such as dust deposition at the ocean surface (Figure 1.7). The regression equation expresses fluxes of POC in the deep ocean (>1500m) as a function of the
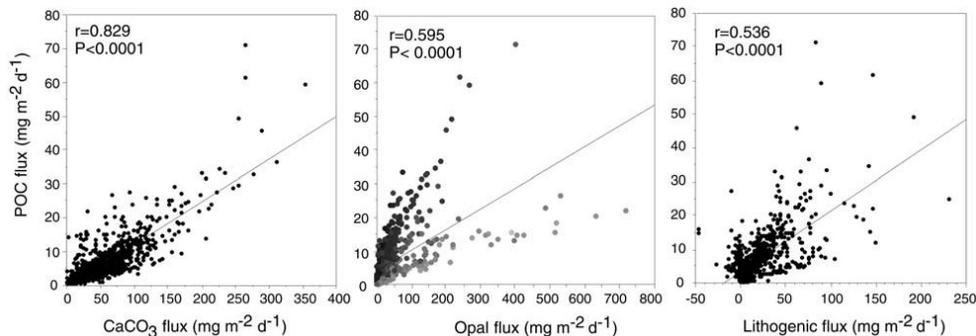
Figure 1.7: Scatterplot of POC fluxes versus $CaCO_3$, opal and lithogenic material fluxes measured by sediment traps below 1000m depth. Correlation coefficients and probability values are given. The slopes of the regression lines are 0.126, 0.061 and 0.180 for $CaCO_3$, opal and lithogenic material respectively. Lighter shading in the central panel corresponds with high opal to total flux ratios. From *Klaas and Archer* (2002).

mineral fluxes related to eachother by 'carrying coefficients' ($\beta_{mineral}$), a statistical parameter relating an amount of POC to each unit of mineral:

$$F_{POC} = F_{CaCO_3} \cdot \beta_{CaCO_3} + F_{opal} \cdot \beta_{opal} + F_{litho.} \cdot \beta_{litho} \qquad (1.10)$$

The ability of the analysis to predict POC fluxes was consistently high ($R^2 \sim 0.95$). The carrying coefficients were relatively large for $CaCO_3$ (β_{CaCO_3}=0.070-0.094), smaller for opal (β_{opal}=0.023-0.030) and more variable for lithogenic material ($\beta_{litho.}$=0.035-0.071) (*Klaas and Archer*, 2002). This analysis suggested that 83% of global POC fluxes are associated with $CaCO_3$. The pattern in carrying coefficients correlates with the density of each mineral type ($CaCO_3$:2.71 g cm^{-3} *cf.* Opal: 2.10 g cm^{-3}; *Gehlen et al.*, 2006) suggesting that the larger density of $CaCO_3$ increases the sinking velocity of particles delivering more POC to depth than other minerals such as opal (*Klaas and Archer*, 2002).

Laboratory experiments where calcifying and non-calcifying coccolithophore cultures are allowed to continuously aggregate and sink in rotating tanks of seawater, have provided evidence that the presence of $CaCO_3$ shells increases the sinking velocity of particles, promotes the formation of compact aggregates (*Ploug et al.*, 2008; *Engel et al.*,

2009a; *Iversen and Ploug*, 2010) and provides some protection against remineralisation (*Engel et al.*, 2009b). However, the strong global quantitative relationship between POC and minerals observed in sediment traps (*Armstrong et al.*, 2002; *Klaas and Archer*, 2002) is not evident in studies of ballasting at smaller spatial and temporal scales. For example, the central panel in Figure 1.7 shows that there are potentially two clusters of data between POC and opal consistent with regional differences in the magnitude of opal export. Consequently, the carrying coefficients and their relative magnitude were significantly different when calculated over ocean regions (*Ragueneau et al.*, 2006; *Lam et al.*, 2011; *Le Moigne et al.*, 2012, 2014) and during different times of the year (*Lutz et al.*, 2007; *Lee et al.*, 2009; *Lam et al.*, 2011). The quantitative association supporting the ballast hypothesis could therefore be an artefact of averaging over different time-scales and spatial scales rather than an explicit mechanism (*De La Rocha and Passow*, 2007; *Armstrong et al.*, 2009; *Lam et al.*, 2011). Chapter 3 presents a novel analysis of the spatial variability in the carrying coefficients calculated from deep sediment trap data to further explore these uncertainties.

The correlations between POC and minerals in sediment trap observations could equally indicate a role for POC in setting mineral fluxes to the deep ocean (*Passow*, 2004). Additional rolling tank experiments have shown that aggregates forming from suspended particles reached 2-5% dry-weight of POC regardless of the amount of minerals available (*Passow and De la Rocha*, 2006; *De La Rocha et al.*, 2008). Particular components of POC may therefore act as a 'glue' for particles aiding aggregation driving the sinking of particles and the statistical relationships in sediment trap observations. For example, Transparent Exopolymer Particles (TEP) have been identified as a significant organic molecule in this role (*Passow*, 2002; *Burd and Jackson*, 2009). Although this mechanism would still link the fluxes of POC and minerals, the associated feedbacks on the carbon cycle may be different.

1.2.2.2   Alternative Mechanisms

A range of alternative mechanisms have also been suggested to be important in understanding the variability in biological pump observations. Alternative explanations of the observed patterns in Figure 1.7 are based on processes driven by processes in surface ocean ecosystems. Differences in the export ratio of POC, *i.e.,* the amount of POC leaving the surface versus productivity, have been shown to be strongly correlated with temperature (*Laws et al.*, 2000; *Henson et al.*, 2011). Low latitude ecosystems where the supply of nutrients via mixing is low are characterised by a high level of organic matter recycling in the surface (*Laws et al.*, 2000). POC leaving the surface may therefore have been already processed many times resulting in relatively refractory POC being exported from the surface that cannot be broken down so easily as it sinks. In contrast, POC produced in the the high latitudes may be relatively more labile, facilitating remineralisation at a much faster rate resulting in shallower flux curves (*Henson et al.*, 2012). The interaction between sinking particles and zooplankton has also been highlighted as important. POC that is consumed by zooplankton and excreted as dense pellets may reach deeper depths and act to homogenise POC and $CaCO_3$ fluxes (*Lam and Bishop*, 2007; *Fischer and Karakas*, 2009; *Alonso-Gonzalez et al.*, 2010; *Lam et al.*, 2011).

Although not directly based on biological pump observations, several studies have also explored the potential for remineralisation rates to be controlled by temperature. Rates of biological processes are sensitive to temperature, approximately doubling for an increase in 10°C (*Eppley*, 1972). Following this logic, all things being equal, an increase (or decrease) in water column temperature could drive an increase (decrease) in remineralisation rates leading to shallower (deeper) flux curves. A significant effect of temperature on organic matter remineralisation has been confirmed from a study incubating samples of seawater at a rnage of temperatures that have been collected from the upper ocean in various locations in the global ocean (*Bendtsen et al.*, 2015). Several

modelling studies have also considered the potential effects of temperature dependent remineralisation (*Matsumoto et al.*, 2007; *Matsumoto*, 2007; *Chikamoto et al.*, 2012; *Segschneider and Bendtsen*, 2013; *John et al.*, 2014). The temperature of the water column also has an effect on the viscosity of seawater that may affect the sinking velocity of particulates. *Taucher et al.* (2014) predict that a increase in water column temperature, leading to a decrease in viscosity, could additionally cause a significant increase the sinking velocity of particles leading to a deeper flux profile.

1.2.3   Implications and Conclusions

As described in section 1.1.3 changes in the biological pump that could potentially occur within the ocean interior, such as changes in the vertical distribution of DIC and the relative amounts of POC and $CaCO_3$ reaching the sediments, could have a large impact on atmospheric CO_2 concentrations. In comparison to changes in the biological pump driven in the surface ocean, changes driven in the ocean interior have received much less attention despite having a comparable effect on atmospheric CO_2. Modern observations of the biological pump in the modern ocean provide a basis for understanding what mechanisms are key to understanding how changes in the ocean interior could have changed atmospheric CO_2 in past climates as well as predicting its role in future climate change. However, the modern observations and the range of mechanistic hypotheses highlighted illustrate that there are still significant uncertainties as to how the biological pump functions in the ocean interior.

The range of mechanisms described above, hypothesised to explain the observations of the modern biological pump in the ocean interior, will respond to a range of different environmental changes in a number of different ways. This has important implications for understanding what role changes in the biological pump in the ocean interior may have had in past changes in atmospheric CO_2, as well as predicting how it may respond to future climate change. For example, by hypothesising that deep fluxes of POC are

mechanistically linked to fluxes of $CaCO_3$ the ballast hypothesis suggests the efficiency of the rain-ratio to alter atmospheric CO_2 (see Section 1.1.3; *Archer and Maier-Reimer*, 1994) may be reduced. *Ridgwell* (2003) suggests that ballasting could reduce the original 70-90ppm reduction of atmospheric CO_2 (from a 40% reduction in the $CaCO_3$ export flux) by 75-80%. Modelling experiments also suggest ballasting may even act to increase atmospheric CO_2 dependent on the extent to which POC and $CaCO_3$ fluxes are coupled (*Barker et al.*, 2003; *Hofmann and Schellnhuber*, 2009).

Alternative interpretations of the observations that support the ballast hypothesis, as well as other hypothesised mechanisms, provide an additional source of uncertainty about the biological pump in the ocean interior. For example, a global increase in the global mean temperature of the ocean may lead to increased remineralisation rates (*Bendtsen et al.*, 2015) but also increased sinking velocities due to the decreased viscosity of seawater (*Taucher et al.*, 2014). These two mechanisms will act to offset eachother in terms of the distribution of DIC as described by POC flux curves. Added to this, temperature is likely an important component in driving ecosystem changes which may also change the characteristics of sinking POC (*Henson et al.*, 2012).

Additionally the uncertainty in mechanistic interpretations is affected by uncertainty in the observations themselves. The relatively small size of sample from regions such as the Southern Ocean (section 1.2.1) may bias the interpretation of ballasting from sediment trap observations (*De La Rocha and Passow*, 2007; *Lam et al.*, 2011). Additionally, the focus on fluxes in the deep ocean may also bias our interpretation of observations (*Ragueneau et al.*, 2006). A recent compilation of sediment trap obsverations from <1000m water depth show a considerably different spatial pattern of POC fluxes (*Marsay et al.*, 2015) (Figure 1.6d). Similar to compilations of deep sea observations, this compilation does not include any observations from the Southern Ocean (*Marsay et al.*, 2015).

Ultimately, due to uncertainties in the interpretation of modern observations of the

biological pump and uncertainties in the observations themselves, our understanding of the biological pump in the ocean interior is still limited. As such, it is still a key uncertainty in our understanding of the carbon cycle and how it functioned in the past and how it may respond in the future (*Passow and De la Rocha*, 2006; *Riebesell et al.*, 2009; *Honjo et al.*, 2014).

1.3   Aims and Outline of the Thesis

The overall aim of this thesis is to evaluate and advance our current mechanistic understanding of the biological pump in the modern ocean interior. A particular focus is made on exploring the global scale function of the biological pump such that the findings are applicable to understanding how they might impact changes in atmospheric CO_2 concentrations. The specific science objectives are:

- **To build on existing analyses to assess and explore uncertainties in the ballast hypothesis to further develop understanding of how this mechanism works:**

    Previous studies have suggested that the multiple linear regression analysis deep sediment trap observations by (*Klaas and Archer*, 2002) (see equation 1.10) does not take into account spatial variability in ballasting relationships. In **chapter 3**, a new updated dataset of global deep sediment trap data is first created. The novel application of Geographically Weighted Regression, that facilitates the carrying coefficients from *Klaas and Archer* (2002) to vary as a function of space, is described. The results are discussed in relation to new mechanistic interpretations of the bioloigcal pump in the ocean interior and the use of ballasting in ocean biogeochemical modelling.

- **To explore and develop new methods of diagnosing high resolution estimates of the spatial variability in remineralisation depths using dissolved tracer observations:**

One of the limitations of using observations such as sediment trap data is that the global sampling density is relatively sparse. **Chapter 4** explores whether model circulation rates can be used to estimate annual average rates of organic matter remineralisation from observed dissolved tracer fields such as $[PO_4]$. Circulation rates are diagnosed from the Earth system model of intermediate complexity GENIE using a novel application of the Transport Matrix Method. The various sources of error when estimating remineralisation rates and possible future research directions are explored.

- **To explore the uncertainties and sensitivity of remineralisation depths using an ocean biogeochemical model:**

  Observed nutrient fields, such as $[PO_4]$, have been used to optimise ocean biogeochemical model parameters and validate mechanistic parameterisations of the biological pump. In **chapter 5**, a new computationally fast phosphorus-only version of GENIE based around the steady-state circulation represented by the Transport Matrix diagnosed in **chapter 4** is described. A Latin hypercube approach is used to sample spatial variability of remineralisation depths in biogeochemical biomes which are then used to simulate steady state $[PO_4]$ fields that can be compared to observed $[PO_4]$. The sensitivity of $[PO_4]$ to regional variability in remineralisation depths is explored before trying to find a set of optimal remineralisation depths that statistically fits observations well. The findings are then compared to the sensitivity of the associated CO_2 uptake from changing remineralisation depths.

CHAPTER 2

Modelling the Biological Pump: A description of the Earth System Model GENIE

Quantitative models of ocean biogeochemistry and climate are an important tool in understanding the dynamics of the carbon cycle as well as providing complementary methods to quantifying processes such as the biological pump in the modern ocean. A range of models have been applied to explore the biological pump, ocean biogeochemistry, and the carbon cycle from box models, which consider the global ocean in terms of a relatively small number (3-12) boxes with fixed ocean transport, *e.g.,* (*Sarmiento and Toggweiler*, 1984; *Matsumoto et al.*, 2002; *Zeebe*, 2012), to 3D-resolution dynamic coupled general circulation models (GCMs) with carbon cycle models. Bridging the gap between these different modelling approaches are Earth system models of intermediate complexity (EMICs): models that include fewer processes than GCMs but include a higher number of interacting components (*Claussen et al.*, 2002). The Grid ENabled Integrated Earth system model (GENIE) is an EMIC that is used in Chapters 4 and 5 of this thesis to explore and quantify the biological pump in the modern ocean. This chapter provides a general background and a description of the components and results that are relevant to the use of GENIE in this thesis.

2.1   Model Description

The version of GENIE used in this thesis is based on the GENIE-1 model (*Lenton et al.*, 2006; *Ridgwell et al.*, 2007a). GENIE-1 comprises of a ocean-atmosphere-sea ice climate model (C-GOLDSTEIN) consisting of a 3-D ocean, a 2-D energy-moisture balance atmosphere model, a dynamic-thermodynamic sea-ice model coupled with a representation of marine biogeochemical cycling (BIOGEM) and a mixed atmospheric chemistry box (ATCHEM) (Figure 2.1). GENIE-1 also includes other components, such as a land surface physics and terrestrial carbon model (ENTS) and a representation of deep-sea sediments (*Ridgwell and Hargreaves*, 2007) which are not used or considered further in this thesis. GENIE-1 was primarily designed for speed of numerical integration to enable the application to palaeo-climate simulations and long-term simulations of the carbon cycle and climate (*Lenton et al.*, 2006).

The configuration of C-GOLDSTEIN in GENIE-1 used in this thesis is non-seasonally forced and implemented on a 36x36 equal-area horizontal grid with grid-boxes that are incremented by 10° in longitude and a uniform increment in the sine of latitude (3.2° latitude at the equator to 19.2° latitude at the poles). The ocean has 8 vertical z-coordinate levels. The horizontal grid and bathymetry are shown in Figure 2.2. Other configurations of C-GOLDSTEIN are also available including a 16 vertical level model with non-seasonal and seasonal forcing that are commonly used with BIOGEM (*Cao et al.*, 2012), as well as higher horizontal and vertical resolutions that have not yet been used with BIOGEM (*Marsh et al.*, 2011). The 36x36x8 configuration is used here as the non-seasonal forcing and availability of alternative circulation states from the original calibration of the circulation parameters facilitate the analysis in Chapter 4. The applicability and limitations of using this configuration are discussed in this chapter as well as in relation to more specific details in Chapters 4 and 5.

The specific version of GENIE-1 used in this thesis is the "muffin" branch of the carbon cycle focused version (cGENIE). The "muffin" branch is a version that contains only
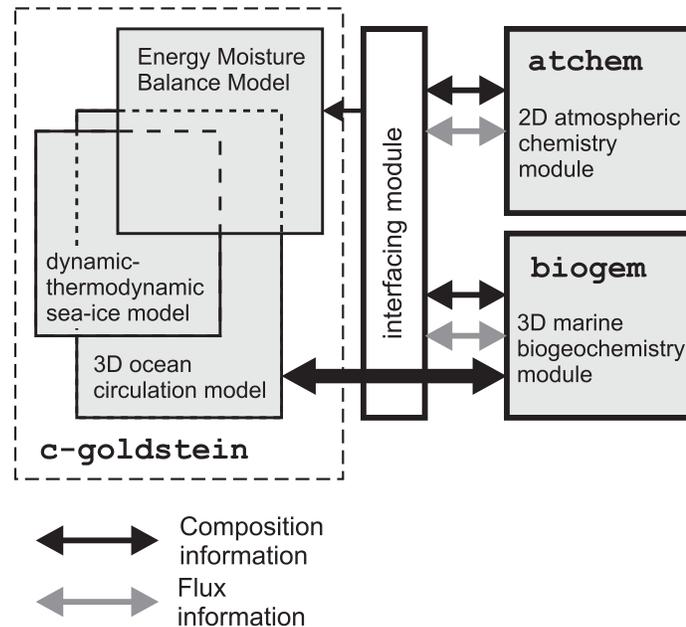
Figure 2.1: Schematic of the relationship between the different components of the GENIE-1 model. Reproduced from *Ridgwell et al.* (2007a).

commonly used components of the original model for studying biogeochemical cycles and the carbon cycle. The code is available from:

`https://svn.ggy.bris.ac.uk/subversion/genie/branches/cgenie.muffin/`

using the username `genie-user` and the password `g3n1e-user`. The svn revision number for the code used in this thesis is 9167. A minor difference between the muffin branch and GENIE-1 used in *Ridgwell et al.* (2007a) is that a default timestep of 96 year^{-1} is used instead of 100 year^{-1}.

2.1.1   C-GOLDSTEIN

The fast climate model, C-GOLDSTEIN, forms the basis for GENIE and is described in full by *Edwards and Marsh* (2005) and more recently by *Marsh et al.* (2011). It features a reduced physics (frictional geostrophic) 3-D ocean circulation based on the thermocline (or plantetary geostrophic) equations with the addition of a linear drag term in the horizontal momentum equations (*Edwards and Shepherd*, 2002; *Edwards and Marsh*, 2005). The resulting circulation is dynamically similar to that of classical
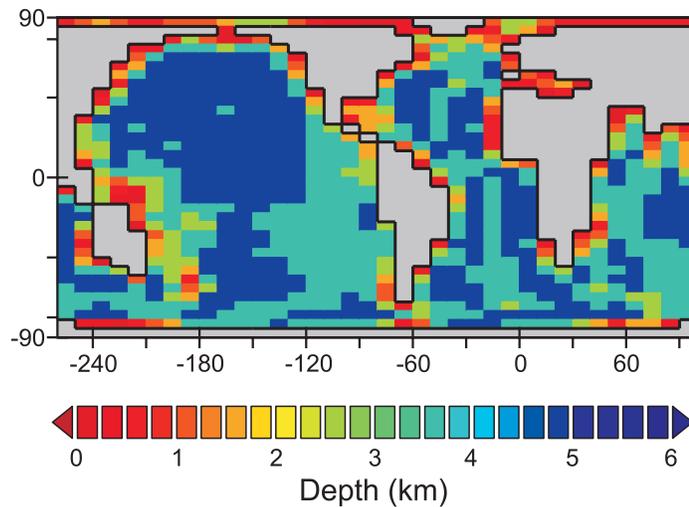
Figure 2.2: The gridded contintental configuration and ocean bathymetry of the 36x36x8 grid of the GENIE-1 model. Reproduced from *Ridgwell et al.* (2007a)

GCMs based on the primitive equations but significantly faster to run. The prognostic variables at each grid-point are temperature, salinity, and a 3-D vorticity field. The model calculates the horizontal and vertical transport of heat, salinity and other tracers (such as the biogeochemical tracers that are transported as passive tracers) through advection, convection and mixing via parameterised iso- and diapycnal diffusion and eddy-induced advection (*Edwards and Marsh*, 2005; *Marsh et al.*, 2011).

C-GOLDSTEIN employs a rigid lid surface boundary condition, *i.e.,* grid cell volumes cannot change in response to evaporation/precipitation or runoff of freshwater from land at the ocean surface. Gains or losses of freshwater are implemented via a virtual salinity flux at the surface. Biogeochemical tracers are salinity normalised prior to the calculation of ocean transport in C-GOLDSTEIN to represent changes in tracer concentrations due to this effect (*Ridgwell et al.*, 2007a).

The energy-moisture balance model (EMBM) of the atmosphere has air temperature and specific humidity as prognostic variables. The model calculates the horizontal transport of heat and moisture in the atmosphere via winds and mixing and precipitation as a funciton of relative humidity (*Marsh et al.*, 2011). The exchange of heat between the atmosphere and the ocean, sea-ice and land is calculated as the balance between

incoming shortwave radiation, sensible and latent heat fluxes, re-emitted long-wave radiation, and outgoing planetary long-wave radiation (*Marsh et al.*, 2011). Exchange of moisture is calculated as the balance between precipitation, evaporation and runoff (*Marsh et al.*, 2011).

The sea-ice model calculates the horizontal transport of sea-ice concentration and thickness, as well as the exchanges of heat and freshwater with the atmosphere and surface ocean (*Marsh et al.*, 2011).

The degree of spatial and temporal abstraction inherent in both C-GOLDSTEIN and BIOGEM (as is common with other models: *Claussen et al.*, 2002) results in parameter values that are not well known and require calibrating by comparing model performance against observations. The parameters for C-GOLSTEIN were calibrated against annual mean climatological observations of temperature, salinity, surface air temperature and humidity using the ensemble Kalman filter (EnKF) methodology (*Hargreaves et al.*, 2004; *Ridgwell et al.*, 2007a), described further in *Annan et al.* (2005). This method has been shown to recover parameter values in identical twin testing using C-GOLSTEIN (*Hargreaves et al.*, 2004; *Annan et al.*, 2005). A useful consequence of the EnKF calibration is that it produces ensemble members that are random samples from the probability distribution function, *i.e.,* plausible estimates of the modern ocean circulation based on different parameter combinations, (*Annan et al.*, 2005; *Ridgwell et al.*, 2007a) that are used in Chapter 4. The parameter set used in this thesis is reported in Table C.2 under SYN.

2.1.2   BIOGEM

The biogeochemical model (BIOGEM) is based on the model by *Ridgwell* (2001) of which the application within the GENIE framework is described by *Ridgwell et al.* (2007a). The following description describes the representation of the biological pump in BIOGEM that is relevant to Chapters 4 and 5 based on the description by *Ridgwell*
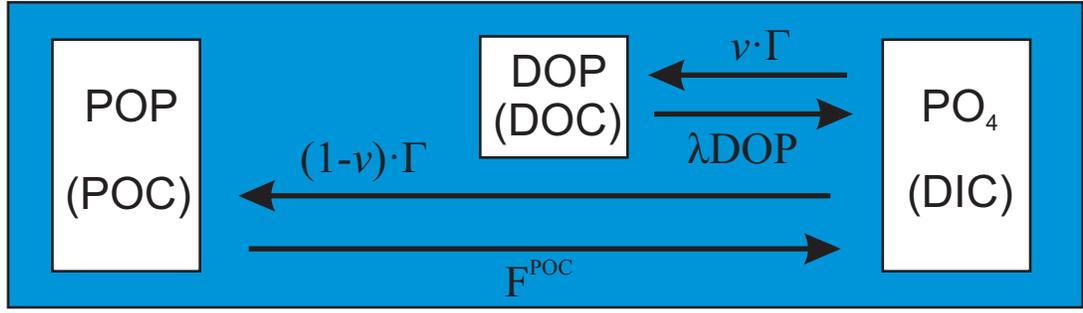
Figure 2.3: Schematic illustrating the soft-tissue pump in BIOGEM. Transformations between the PO_4, dissolved organic phosphorous (DOP) and particulate organic phosphorous (POP) are governed by equations 2.1, 2.2, 2.3 and 2.4. Carbon equivalents are given in brackets. Figure adapted from *DeVries et al.* (2014).

et al. (2007a).

BIOGEM estimates export production based on limitation by a single nutrient, phosphate (PO_4). The biological uptake of PO_4 is then related to other tracers, such as oxygen (O_2), nitrogen and carbon, via global average molar ratios of phytoplankton nutrient uptake of -170:1 (*Anderson and Sarmiento*, 1994), 16:1 (*Redfield et al.*, 1963), and 106:1 (*Redfield et al.*, 1963). The governing equations (see also Figure 2.3 for a schematic illustration) for PO_4 and dissolved organic phosphorous (DOP) in the surface ocean omitting the ocean transport terms are:

$$\frac{\delta \mathrm{PO_4}}{\delta t} = -\Gamma + \lambda \mathrm{DOP} \tag{2.1}$$

$$\frac{\delta \mathrm{DOP}}{\delta t} = v\Gamma - \lambda \mathrm{DOP} \tag{2.2}$$

where Γ is the biological uptake of PO_4 which is calculated from an assumed maximum uptake rate $u_0^{\mathrm{PO_4}}$ (mol PO_4 kg^{-1} yr^{-1}), that occurs in the absence of any limitation on phytoplankton growth; a Michaelis-Menten type kinetic limitation of nutrient uptake where $K^{\mathrm{PO_4}}$ is the half-saturation constant; A is the fractional sea-ice coverage calculated by the sea-ice model; and I is the the strength of solar insolation normalised by the solar constant (I_0):

$$\Gamma = u_0^{\text{PO}_4} \cdot \frac{\text{PO}_4}{\text{PO}_4 + K^{\text{PO}_4}} \cdot (1 - A) \cdot \frac{I}{I_0} \tag{2.3}$$

A proportion (v) of the PO_4 taken up by biology in Equation 2.1 is partitioned into DOP which is remineralised with a time constant of $\frac{1}{\lambda}$. The remaining PO_4 is partitioned into a particulate organic phosphorous fraction that is exported vertically from the surface grid-box (hereafter this is related to particulate organic carbon, POC, via the Redfield ratio stated above). The change in the POC flux with depth (F_z^{POC}) in the water column occurs instantaneously and is represented using the sum of two exponential decay functions:

$$F_z^{\text{POC}} = F_{z=h_e}^{\text{POC}} \cdot \left((1 - r^{\text{POC}}) \cdot \exp(\frac{zh_e - z}{l^{\text{POC}}}) \right)$$

$$+ F_{z=h_e}^{\text{POC}} \cdot \left(r^{\text{POC}} \cdot \exp(\frac{zh_e - z}{l^{\text{rPOC}}}) \right) \tag{2.4}$$

where r^{POC} is an initial partitioning of POC between labile and refractory components at the surface which decay with depth according to length scales (l^{POC}) and (l^{rPOC}) respectively. The refractory length scale (l^{rPOC}) is set to $1\text{x}10^5$m such that almost all refractory POC reaches the deep ocean. The choice of a double exponential formulation (*Andersson et al.*, 2004) over a power-law curve, *e.g.,* the 'Martin Curve' (*Martin et al.*, 1987), reflect that they produced a reasonable interior ocean [PO_4] distribution (see Figure 2.6 and discussion in following section) due to the refractory component whilst retaining an approximation to the Martin Curve at shallower depths (*Ridgwell*, 2001; *Cameron et al.*, 2005). All POC reaching the deepest ocean grid-box is remineralised fully back to PO_4 such that there is no net gain or loss from the ocean, *i.e.,* it is a closed system (although note that GENIE can include a sediment module (SEDGEM: *Ridgwell and Hargreaves*, 2007).

Although not used in the thesis but considered here for completeness, the export flux

of calcium carbonate ($F_{z=h_e}^{CaCO_3}$) from the surface, precipitated at the surface by biology, is related to the particulate organic carbon flux exported from the surface grid-boxes ($F_{z=h_e}^{POC}$) via:

$$F_{z=h_e}^{\text{CaCO}_3} = \gamma \cdot r_0^{\text{CaCO}_3:\text{POC}} \cdot F_{z=h_e}^{\text{POC}} \tag{2.5}$$

where $r_0^{\text{CaCO}_3:\text{POC}}$ is a spatially uniform scalar and γ is a thermodynamically-based local modifier of the rate of carbonate production based on the local saturation state with respect to calcite (Ω) and a constant (η):

$$\gamma = (\Omega - 1)^\eta \text{ if } \Omega > 1.0$$

$$= 0.0 \text{ otherwise} \tag{2.6}$$

The dissolution of $CaCO_3$ in the water column is treated in the same manner as POC (Equation 2.4).

As with C-GOLDSTEIN, parameters in BIOGEM were calibrated using the EnKF methodology against annual mean climatological observations of phosphate and alkalinity (*Ridgwell et al.*, 2007a; *Ridgwell and Hargreaves*, 2007; *Cao et al.*, 2012) using the ocean circulation and climate previously calibrated by *Hargreaves et al.* (2004). Parameters for the BIOGEM calibration are described in Table 5.1 and *Ridgwell and Hargreaves* (2007). Note that this parameter set is slightly different to that originally reported by *Ridgwell et al.* (2007a).

2.2   Model Evaluation

This section describes results for the pre-industrial model state derived from the 36x36x8 non-seasonally forced configuration corresponding to the model state used in Chapter

4 (see experiment SYN but note this simulation is performed with a Martin curve rather than the exponential formulation of POC in equation 2.4) and the inital spin-up used in Chapter 5. The description here is based on the more detailed descriptions in *Hargreaves et al.* (2004), *Edwards and Marsh* (2005) for C-GOLDSTEIN and *Ridgwell et al.* (2007a) for BIOGEM. The results shown here are kept relevant to the scope of this thesis: a focus on the ability of C-GOLDSTEIN to simulate large-scale circulation patterns and the ability of BIOGEM to simulate large-scale patterns in observed [PO$_4$].

2.2.1   Circulation

The large-scale circulation is shown by the overturning circulation in GENIE in Figure 2.4. The model simulates an Atlantic meridional overturning circulation (AMOC) with an upper cell transporting around 16 Sv and an abyssal cell transporting around 0.2 Sv (Figure 2.4a) consistent with observations of the AMOC of 15 Sv and 2 Sv respectively (*Ganachaud and Wunsch*, 2000). The global overturning has a maximum of 40 Sv and a minimum of -44 Sv (Figure 2.4c). *Hargreaves et al.* (2004) show that the overturning circulation and heat transport in this version of GENIE are within the range of found in models used for the Coupled Model Intercomparison Project (CMIP).

The low resolution and simplified physics of C-GOLDSTEIN confer some limitations on the use of the model, particularly in comparison with higher resolution and more complex GCMs. For example, the low horizontal resolution means that smaller scale features of the ocean circulation are not resolved, *e.g.,* the Indonesian throughflow and Agulhas leakage are simulated in GENIE (*Hargreaves et al.*, 2004). This will limit the study of biogeochemistry at these scales, although this is still an issue for higher resolution models, *e.g., DeVries et al.* (2014) removed [PO$_4$] in the North West Pacific from the comparison against observations when modelling remineralisation as the circulation was not well constrained in this area in their model. The horizontal resolution is also an issue for sea-ice in the Southern Ocean. The area in which sea-ice
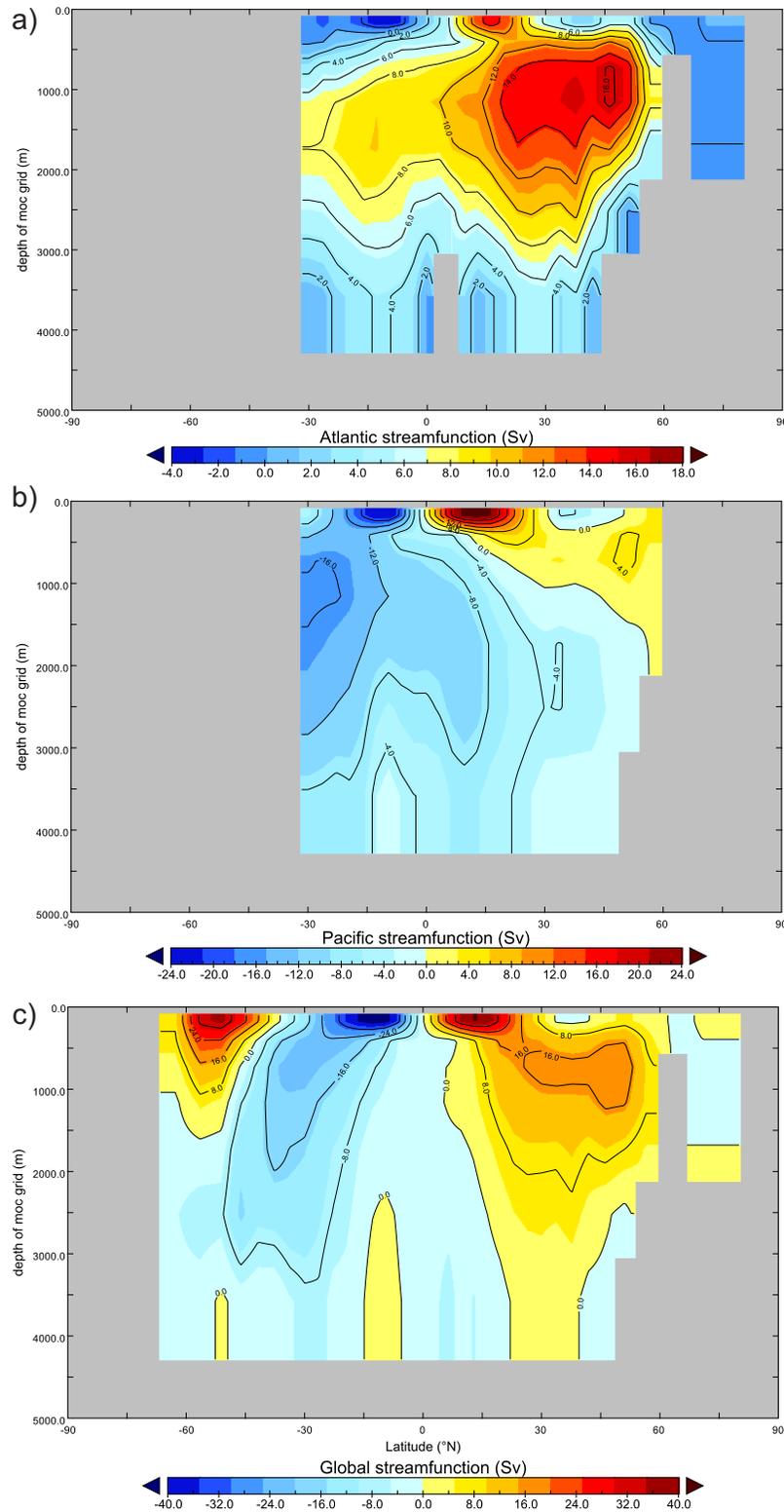
Figure 2.4: Zonally averaged overturning circulation (Sv) in each basin simulated by GENIE. Positive values indicate a clockwise circulation cell and negative values indicate an anti-clockwise circulation. The magnitude of the values indicates the strength of the circulation.

is expected falls within the southern-most grid-boxes around Antarctica (*Marsh et al.*, 2011) leading to sea-ice not being simulated in the Southern Ocean in this version of GENIE. This will have some impact on the simulated export production as this is in part limited by sea-ice (equation 2.3). Overall, the large-scale circulation patterns are well simulated (*Hargreaves et al.*, 2004).

The low vertical resolution of GENIE (8 z-coordinate levels in the vertical) leads to a less stable water column at high latitudes resulting in excessive ventilation (*Müller et al.*, 2006; *Ridgwell et al.*, 2007a). This is demonstrated in *Ridgwell et al.* (2007a) by forcing the model with the observed historical trajectory of atmospheric CO_2 between 1765 to 2000. At the year 1994, the anthropogenic CO_2 inventory is 171 PgC compared with a data-based estimate of 118$\pm$19 GtC (*Sabine et al.*, 2004; *Ridgwell et al.*, 2007a). Comparison of the water column integrated anthropogenic CO_2 shows that the largest deviations from data-based estimates are in the Southern Ocean and eastern North Pacific (see Figure 2.5). This may have implications for biogeochemical tracers. However, this effect is significant on decadal timescales and is less important on timescales of >100 years (*Ridgwell et al.*, 2007a).

2.2.2   Biogeochemistry

The zonal average fields of $[PO_4]$ for each basin in GENIE (Figure 2.6) agree with observations of $[PO_4]$ to a first-order. For example, the model captures the high concentrations in the deep North Pacific (2.6), the low values of $[PO_4]$ associated with deep water formation in the North Atlantic (Figure 2.6). However, GENIE predicts higher concentrations than observed in the Atlantic sector of the Southern Ocean (Figure 2.6b *cf.* 2.6a) which is indicative of the excess ventilation noted above (*Ridgwell et al.*, 2007a). Surface concentrations of $[PO_4]$ are over-estimated in the West Equatorial Pacific and Equatorial Indian Ocean compared to observations (Figure 2.6h *cf.* 2.6g). This occurs due to over-estimated upwelling in these regions (*Ridgwell et al.*, 2007a). The
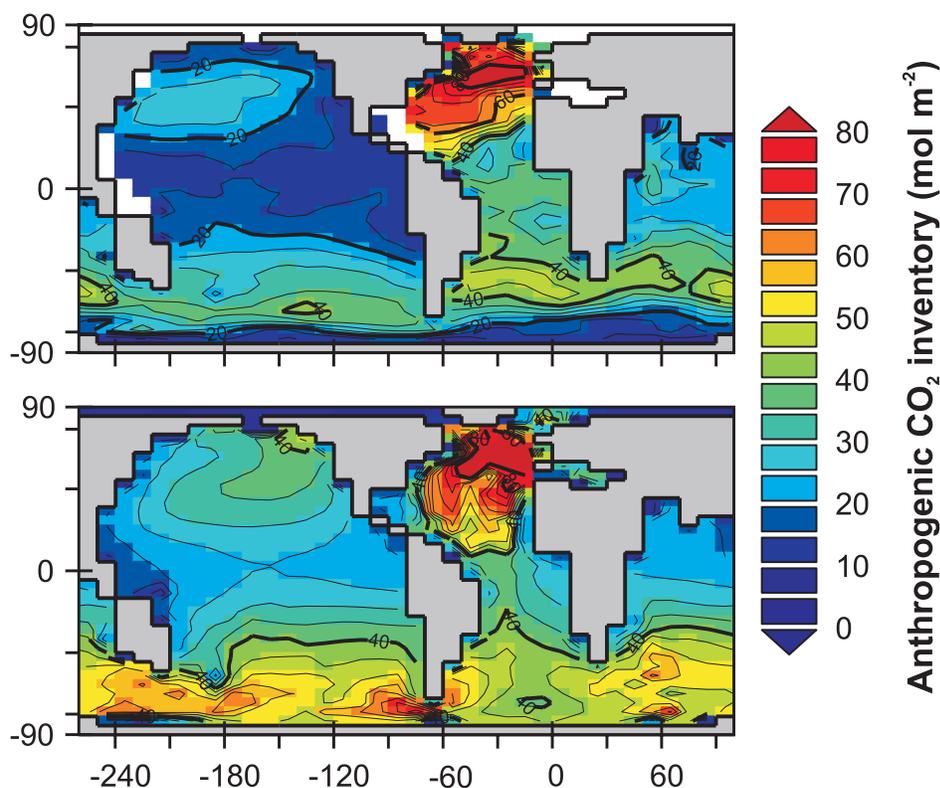
Figure 2.5: Water column integrated anthropogenic CO_2 inventory in the ocean for the year 1994 for (top) observed (*Key et al.*, 2004) and (bottom) modelled. Adapted from *Ridgwell et al.* (2007a).

global export of POC is 9.01 PgC year^{-1} which is consistent with data-based estimates of 5-10 PgC year^{-1} (*Henson et al.*, 2011). The close match to observations of global POC export suggests that the excess ventilation is unlikely to bias the biogeochemical cycling in GENIE (*Ridgwell et al.*, 2007a).

As with C-GOLDSTEIN, the model structure of BIOGEM confers limits on its use in comparison to other biogeochemical models. The model structure is similar to higher resolution biogeochemical models used to quantify the biological pump such as used by *Kwon and Primeau* (2006), *Kriest et al.* (2010), *Kriest et al.* (2012), *Kwon et al.* (2009) and *DeVries et al.* (2014). These models generally predict the magnitude of export by relaxing surface concentrations of [PO$_4$] to observations (*Najjar et al.*, 2007), have a representation of DOM and model POC fluxes using a Martin Curve or equivalent exponential function. The Michaelis-Menten type uptake used in GENIE is therefore dynamically more representative however the single nutrient limitation neglects addi-
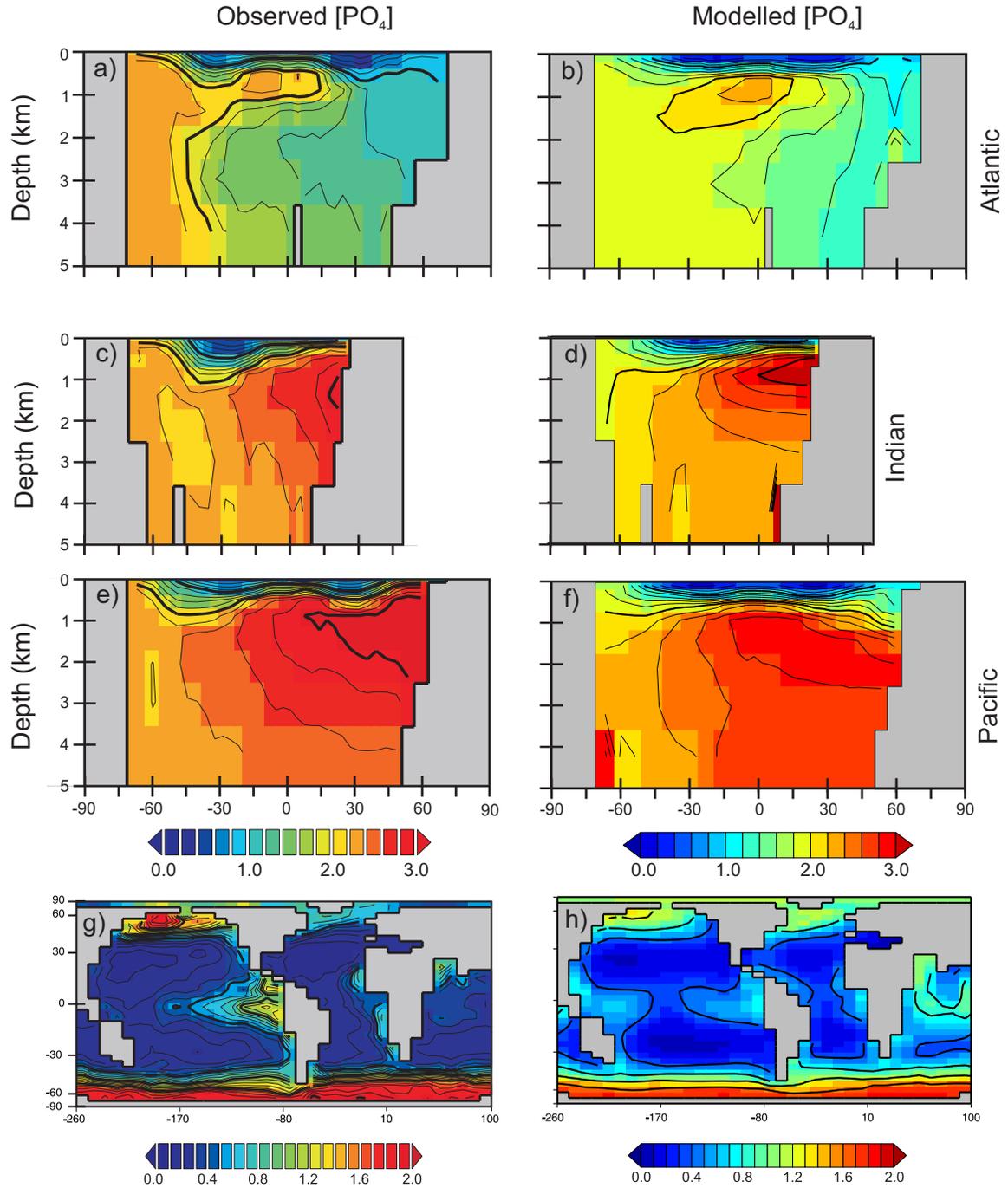
Figure 2.6: Zonal averages of [PO$_4$] for each basin simulated by GENIE (panels b, d, and f) compared against observations (a, c, and e) from *Key et al.* (2004) as shown in *Ridgwell et al.* (2007a). The surface ocean (upper-most grid-box) [PO$_4$] simulated by GENIE is also shown in panel h compared against observations integrated over the top 75m in panel g. Note the small difference in colour scale at low values between model results and observations due to model results being replotted with a slightly different parameter set to *Ridgwell et al.* (2007a). Figure is adapted from *Ridgwell et al.* (2007a).

tional limitations such as by iron (*Martin*, 1990; *Ziegler et al.*, 2013). *Ridgwell et al.* (2007a) notes that the single global parameters controlling export may therefore compensate for utilising most [PO$_4$] in the low-latitudes but not utilising too much in the high nutrient low chloropyll (HNLC) regions. As such surface [PO$_4$] is generally slightly lower in HNLC regions (Figure 2.6h *cf.* g: *Ridgwell et al.*, 2007a). The biogeochemical model also does not simulate different types of phtoplankton or their interactions with zooplankton, such as done in more complex biogeochemical models like MEDUSA (*Yool et al.*, 2013), which have been highlighted as potentially important in driving changes in the biological pump (*Lam and Bishop*, 2007; *Fischer and Karakas*, 2009; *Henson et al.*, 2012; *Lam et al.*, 2011). However, *Kriest et al.* (2010) showed that tracers such as [PO$_4$] were sensitive to only a few key parameters such as remineralisation depth and export. Therefore, relatively simple models similar to the complexity of GENIE were found to be able to reproduce tracer fields to a similar degree as more complex models (*Kriest et al.*, 2010).

2.3   Conclusions

GENIE is a relatively low resolution EMIC designed for fast computation speeds which aid the application to long timescales. The model can reproduce large-scale circulation patterns and observed patterns in [PO$_4$] although the low resolution lultimately limits the ability of the model to reliably simulate smaller scale features. Despite this, there are a number of advantages to the use of GENIE to study the biological pump. Current observations of POC fluxes in the ocean interior suggest there are significant spatial differences in the depth of remineralisation although the nature of this variability is still very uncertain, *e.g., Henson et al.* (2012) *cf. Marsay et al.* (2015). In light of this the ability of GENIE to reproduce large-scales is more than adequate and potentially preferable as more complex models may add additional uncertainty and degrees of freedom. The speed of GENIE also facilitates a modelling approach seeking to quantify the

uncertainty, similar to the approaches used by previous studies (*Kwon and Primeau*, 2006; *Kriest et al.*, 2012). Lastly, modelling the biological pump in GENIE can be consistently applied to the additional uncertainty about mechanisms over long timescales (>1000-10 000 years) and its role in palaeoclimate.

CHAPTER 3

Spatial variability in particulate organic matter and mineral sinking fluxes in the ocean interior

3.1   Introduction

Sinking particles transfer particulate organic carbon (POC) and associated nutrients from the upper ocean to the deep ocean and sediments in a process known as the biological pump (*Honjo et al.*, 2008). These particles, ultimately derived from the growth of phytoplankton at the sunlit surface and carbon fixation through photosynthesis, are also often associated with biominerals such as biogenic silica (opal) and calcium carbonate ($CaCO_3$). As these particles sink, the majority of POC and associated nutrients are remineralized (predominantly) by bacterial metabolic processes and zooplankton flux feeding in the upper $\sim$1000m, leaving a small (5-10%) fraction (relative to that at $\sim$100m) sinking to depth (*Stemmann et al.*, 2004; *Loubere et al.*, 2007; *Honjo et al.*, 2008). Understanding the processes that control the efficiency of the biological pump in transporting carbon and nutrients to depth is key to understanding how the marine carbon cycle functions and regulates atmospheric carbon dioxide (CO_2), *e.g., Archer and Maier-Reimer* (1994).

The ratio of particulate inorganic carbon (PIC) to POC within sinking particles is

known as the 'rain ratio' and is potentially a key quantity that could drive changes in the carbon cycle via the deep ocean and sediments. For instance, on time-scales of a few thousand years, an initial reduction in the PIC flux exported from the surface such that the rain ratio is reduced by 40% could, in theory, lead to a 70-90ppm drawdown of atmospheric CO_2. This drawdown is driven by the remineralisation of an increased POC content in the sediments relative to PIC, lowering the saturation state of the sediment porewater leading to increased carbonate dissolution and therefore an increase in the ocean alkalinity inventory (*Archer and Maier-Reimer*, 1994) (See also discussion in Section 1.1. However, studies based on the analysis of deep sediment trap data in the modern ocean have observed a strong global correlation between mass fluxes of POC and $CaCO_3$, suggesting some mechanism of coupling exists between these important parameters at depth (*Armstrong et al.*, 2002). The resulting 'ballast hypothesis' posits that $CaCO_3$, and to a lesser extent opal and lithogenic material, facilitate the flux of POC to the deep ocean. The presence of minerals in particles may increase the mean density of particles resulting in associated POC sinking through the water column at a faster rate or that POC may be protected from degradation by minerals (*Armstrong et al.*, 2002; *Klaas and Archer*, 2002). If true, this would have the effect of buffering changes in the rain ratio originating at the surface and reducing the potential for altering atmospheric CO_2 (*Ridgwell*, 2003). In the context of ocean acidification and the potential for decreased pelagic calcification driven by falling surface ocean carbonate saturation, ballasting creates a positive feedback to CO_2 by reducing the efficiency of the biological pump, *i.e.,* similar changes in the rain ratio as posited by *Archer and Maier-Reimer* (1994) could lead to an increase in atmospheric CO_2 (*Barker et al.*, 2003; *Heinze*, 2004; *Riebesell et al.*, 2009).

The ballast hypothesis is based on a statistically significant global linear relationship between POC and $CaCO_3$. In response, a range of laboratory and field studies have been conducted to validate the hypothesised mechanisms. Particles have been shown to sink faster under laboratory conditions due to the relatively high density of $CaCO_3$

(*Ploug et al.*, 2008; *Engel et al.*, 2009a; *Iversen and Ploug*, 2010), supporting the ballast hypothesis. However, other lines of evidence point to alternative interpretations. For instance, rolling tank experiments showed the POC:mineral ratio was dependent on the amount of POC present acting as a 'glue', aggregating mineral fluxes (*Passow*, 2004; *Passow and De la Rocha*, 2006; *De La Rocha et al.*, 2008) and controlling the sinking of $CaCO_3$ rather than *vice versa*. Alternatively, variations in surface ecosystem composition might dictate the packaging and remineralization of particles (*Francois et al.*, 2002; *Lam and Bishop*, 2007; *Lam et al.*, 2011) and give rise to the correlations between POC and minerals observed in the deep ocean. These alternative explanations significantly challenge our mechanistic understanding of the dynamics of POC fluxes and create substantial uncertainty in both the magnitude and sign of carbon cycle feedbacks to possible future perturbations (*Barker et al.*, 2003; *Riebesell et al.*, 2009).

The global sediment trap analysis of *Klaas and Archer* (2002) has been highly influential in quantifying the correlation between POC and $CaCO_3$ and helping to formulate the ballasting hypothesis. In that study, the mass flux of POC was expressed as a linear function of three dominant mineral fluxes ($CaCO_3$, opal and lithogenic material), using multiple linear regression analysis (MLRA: Equation 1.10). The derived 'carrying coefficients' (the regression coefficients) were largest for $CaCO_3$ (0.070-0.094), lowest for opal (0.023-0.030), and rather variable for lithogenics (0.035-0.071) (*Klaas and Archer*, 2002) with the resulting statistical models able to explain a large proportion of the observed variability in POC flux. The three mineral model provides a basis for understanding global variability in POC to mineral ratios and can replace this term in the mechanistic model of *Armstrong et al.* (2002), making this a useful method for parameterizing particle fluxes in a range of ocean carbon cycle models (*Howard et al.*, 2006; *Oka et al.*, 2008; *Hofmann and Schellnhuber*, 2009; *Yool et al.*, 2013).

The underlying assumption in analysing the global database in this way is that the statistical relationships (the coefficient values) are the same for any location in the ocean, *i.e.,* they are assumed stationary in space. The use of these global statistical

relationships in models then explicitly makes this same assumption. However, it is reasonable to expect that these relationships may not be constant in space (or time), *i.e.,* they may exhibit spatial nonstationarity, which can be characterised by a non-random distribution of residuals in space (*Fotheringham et al.*, 1998). This potentially raises issues for the interpretation of global regression coefficients and their explicit use as a parameterisation in modelling studies (*Howard et al.*, 2006).

Spatial variability in the relationship between POC and minerals was noted in sediment trap data by *Passow and De la Rocha* (2006) and *Ragueneau et al.* (2006) who suggested global MLRA was, therefore, inappropriate and may have misleadingly resulted in the low carrying coefficients obtained for opal and lithogenics. *Boyd and Trull* (2007) also note that using global annual mean fluxes ignores a large part of variability resulting from processes like El Niño and the biogeochemical sources of the fluxes. A previous basin-scale analysis (Table 3.1) showed considerable regional variability in the dominance of one mineral over another. Global MLRA may then be hiding important regional variability which has implications for how the ballast mechanism is interpreted and particularly for how it is mechanistically implemented in global models.

The spatial variability of the carrying coefficients of ballast minerals is important for understanding the previously observed differences between global coefficients and those seen from individual sites (Table 3.1). *Ragueneau et al.* (2006) took the first step in this respect and applied MLRA to sediment trap data divided by major ocean basin. This broad delineation was reasonably justified but further reduction of the spatial scale poses particular problems. Smaller spatial groupings for regression could be justified, such as biogeochemical provinces (see *Vichi et al.*, 2011) but this introduces a level of subjectivity, as well as problems with the relatively sparse sampling coverage of sediment trap datasets compared to the number of biogeographical provinces.

Table 3.1: Carrying coefficients, *i.e.*, *Klaas and Archer* (2002), derived from previous multiple linear regression analyses applied to a range of global, regional and time-series sediment trap datasets. Ranges given indicate 95% range of carrying coefficients (2x standard error). Adapted from Boyd and Trull (2007).

Global Annual Average		**CaCO$_3$**	**Opal**	**Lithogenic**[c]	**R^2**
Klaas and Archer (2002)[d]	n=78	0.075	0.029	0.052	-
	>2000m	0.064-0.086	0.020-0.037	0.034-0.070	
Francois et al. (2002)[d]	n=62	0.074	0.015	0.074	0.93
	>2000m	0.064-0.084	0.08-0.022	0.051-0.097	
Ragueneau et al. (2006)	n=189	0.081	0.031	0.035	0.89
	>1000m	0.073-0.089	0.023-0.039	0.029-0.041	
Regional Annual Average					
Ragueneau et al. (2006)	n=84	0.077	0.171	0.031	0.87
(Atlantic)	>1000m	0.053-0.101	0.047-0.186	0.023-0.039	
Ragueneau et al. (2006)	n=16	0.026[a]	0.201	0.015[a]	0.96
(Indian)	>1000m	-0.007-0.057	0.123-0.279	-0.079-0.049	
Ragueneau et al. (2006)	n-89	0.063	0.041	0.024	0.95
(Pacific)	>1000m	0.055-0.071	0.035-0.046	0.018-0.030	
Regional Time-Series					
Wong et al. (1999)[d]	1982-1993	0.021	0.013	0.0233	0.69
(Ocean Station P)	3800m	0.002-0.039	0.008-0.034	0.170-0.297	
Conte et al. (2001)[d]	1978-1984	0.045	0.063	0.065	0.98
(Bermuda SCIFF)	3200m	0.038-0.053	0.024-0.102	0.034-0.096	
Honda and Watanabe (2010)[b]	1998-2006	0.025	0.044	-0.006	0.92
(W.Pacific Subarctic Gyre)	4810	n/a	n/a	n/a	

[a] Values from multiple regression analysis were insignificant at $p > 0.05$.

[b] Data was normalized to average of each time-series component before regression analysis

[c] All lithogenic material is estimated as Total Mass-(POCxPOM Conversion factor)+CaCO$_3$+Opal, except *Honda and Watanabe* (2010) which was derived from Al measurements

[d] Adapted from Table 3a in *Boyd and Trull* (2007)

3.1.1   Aims

In this chapter, the spatial variability in ballasting carrying coefficients is explored using an updated sediment trap dataset as a further test of the ballast hypothesis and to further explore the controls on POC fluxes in the modern ocean. To achieve this, the chapter describes the novel application of Geographically Weighted Regression (GWR), that allows coefficients to vary in space and helps avoid the problems stated above. The dataset and technique are described in section 3.2 and are applied using the carrying coefficient approach of *Klaas and Archer* (2002) to explore the spatial variability of the these statistical parameters.

3.2   Methods

3.2.1   Sediment Trap Data

An updated global sediment trap dataset has been collated for the analysis in this chapter. The majority ($\sim$85%) of the dataset used here is from the U.S. Joint Global Ocean Flux Study (JGOFS) available online via: http://usjgofs.whoi.edu/mzweb/syn-mod.htm. The JGOFS dataset provides a synthesis of sediment trap data from a range of study designs including deployments through time both on inter and intra-annual scales and deployments across various transects. The dataset also has broadly consistent methodology across the different studies. A full description of the JGOFS dataset and methodologies can be found in *Honjo et al.* (2008). Other published datasets, comprising the remaining $\sim$15% of the dataset used here, were obtained from additional studies and the World Data Centre for Environmental Sciences (WDC-MARE) online database. The flux and metadata from sediment traps used in this chapter are available in Appendix A.

A range of review criteria were used to select sediment trap data in relation to a number

of potential biases:

- *Sampling Period:* The selected data includes only sediment traps that sampled over a minimum period of 320 days to maximize the quantity of data and its spatial coverage whilst retaining a reasonable annual coverage, *i.e.,* the sampling period covers 88% of a year such that a the period covers at least 3 and half seasons for areas that may significant seasonal variability in fluxes.

- *Available Data:* Data were excluded if any observations were missing for the major components required for the analysis (total mass flux, POC, PIC, biogenic silica).

- *Sampling Depths - Trapping Efficiencies:* Sediment traps at relatively shallow water depths (approx. <1000-1500m) have been shown to be inefficient at trapping particulate material due to the effect of water flow over the trap entrance (*Scholten et al.*, 2001; *Yu et al.*, 2001; *Buesseler et al.*, 2007). At shallow depths, the presence of zooplankton are also an additional uncertainty (*Buesseler et al.*, 2007). For this reason, and to be consistent with the bulk of previous work on the ballast hypothesis, only flux data at >1500m water depth are considered here. Approximately 90 trap samples from the JGOFS dataset are shallower than 1500m when all other criteria are considered. Previous studies have used data from shallower sediment traps but corrected the flux data for an estimated minimum 40% undersampling bias based on sediment trap calibration studies (*Lutz et al.*, 2002, 2007). This approach was not used here as any spatial variability in sampling biases cannot be accounted for and may introduce unknown biases into any spatial analysis.

- *Sampling Depths - Sediments:* The majority of the sediment trap data are >200m above the seafloor (∼97%). 20 traps had unreported bottom depths. Previous studies have considered the potential effect of resuspended sediments as a bias for sediment traps (*Lutz et al.*, 2002, 2007) using 200m above the sediment as a criteria. This is not used here as the dataset is mostly 200m above the sediments

and ballast hypothesis studies (*Klaas and Archer*, 2002; *Ragueneau et al.*, 2006) have not considered this bias.

The mass fluxes of $CaCO_3$ and opal (as g m^{-2} year^{-1}) were estimated from PIC and biogenic Silica (bSi) using mass ratio conversion factors of 8.33 and 2.14 respectively (*Mortlock and Froelich*, 1989). These mass ratios are the same as used in the original JGOFS dataset:

$$CaCO_3 = PIC * 8.33 \tag{3.1}$$

$$Opal = bSi * 2.14 \tag{3.2}$$

The flux of lithogenic material was estimated as the remaining fraction of total mass flux when $CaCO_3$, opal and particulate organic matter (POM) are subtracted. POM was calculated using a coefficient of 1.86. In a small number of cases this produced negative values for lithogenic flux, which were then treated as zero as in *Salter et al.* (2010):

$$POM = POC * 1.86 \tag{3.3}$$

$$Litho. = Total\ Mass - (POM + CaCO_3 + Opal)\ if\ Litho. > 0$$
$$= 0\ otherwise \tag{3.4}$$

The resulting dataset comprises 156 individual sediment trap observations which include data on POC, PIC (as $CaCO_3$), biogenic silica (as opal), lithogenic and total mass fluxes. The dataset includes observations from 25 biogeochemical provinces of the 56 defined by
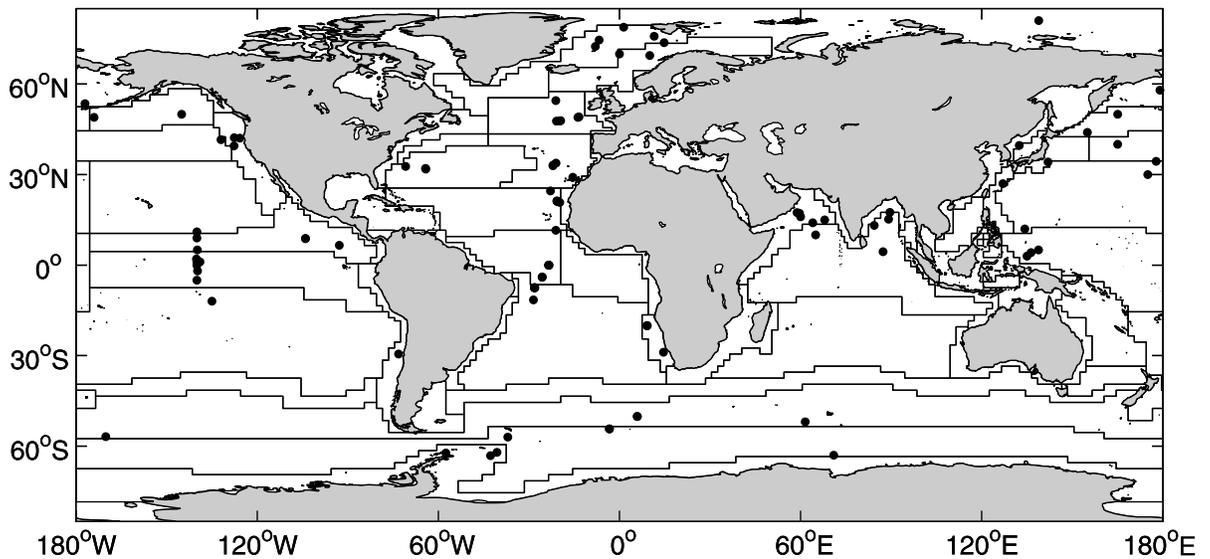
Figure 3.1: Locations of global sediment traps at >1500m. Data locations are given in the context of biogeochemical provinces as per *Longhurst* (1998).

Longhurst (1998) (Figure 3.1). In comparison to the previous global datasets (*Francois et al.*, 2002; *Klaas and Archer*, 2002), this dataset is larger in size (n=156 *cf.* n=62-78) and provides greater spatial coverage, particularly for the southern hemisphere. The dataset is of a comparable size to the recent sediment trap collation by *Honjo et al.* (2008). Approximately half of the dataset is in common with previous analyses, so it is expected *a priori* that the results of a global analysis will not be substantially different from previous studies.

3.2.2 Regression Analysis

3.2.2.1 Global Regression Model

The multiple linear regression analysis used in *Klaas and Archer* (2002) is applied to the updated global sediment trap data. The basic global regression analysis expresses the flux of POC (F_{POC}) in the deep ocean (here defined as >1500m) as a function of the fluxes of CaCO$_3$ (F_{CaCO_3}), Opal (F_{Opal}) and lithogenic material (F_{litho}) at depth:

$$F_{POC} = \beta_0 + \beta_{CaCO_3} \cdot F_{CaCO_3} + \beta_{Opal} \cdot F_{Opal} + \beta_{litho} \cdot F_{litho} \tag{3.5}$$

The regression coefficients (β) in Equation 3.5 relate the strength of the relationship between POC and each mineral type. In the global regression of *Klaas and Archer* (2002), the coefficient values are a global parameter. Previous analyses assumed that the regression passed through the origin (*Ragueneau et al.*, 2006) requiring that the flux of POC must be zero when the flux of minerals is zero. An additional intercept term (β_0) is included here so that the analysis is amenable and directly comparable to the geographically weighted technique employed and also to create a more general model in which it is possible that there could be additional POC not directly associated with the mineral flux. This can be related directly to the model of *Armstrong et al.* (2002) (Equation 1.9) where Equation 3.5 without the intercept term is the $F_{POC}(\infty)$ term in Equation 1.9, *i.e.,* all POC at depth that is associated with mineral fluxes. The intercept term allows for some residual POC flux not associated with the mineral fluxes.

3.2.2.2   *Geographically Weighted Regression Model*

Geographically weighted regression is a relatively novel but simple technique of regression which allows the estimation of local statistical parameters. Ror a full description see *Fotheringham et al.* (2002), also *Brunsdon et al.* (1998) and *Fotheringham et al.* (1998). The global multiple linear regression model from (3.5) can be considered more generally as:

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \epsilon_i \tag{3.6}$$

where k predictors (x) are used to predict y at the i_{th} point in space. The global model can be re-written to estimate local parameters as:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \epsilon_i \tag{3.7}$$

where (u_i, v_i) denotes the coordinates of the ith point in space. $\beta_k(u_i, v_i)$ is then a realisation of the continuous function $\beta_k(u, v)$ at point i (*Fotheringham et al.*, 1998). The global model can be seen as a case where the parameter surface is considered constant in space. GWR approximates the above equation by selecting a subset of data around i which is weighted according to their distance from i. This assumes that parameters display a degree of spatial consistency, *i.e.,* that parameters become increasingly different as distance increases from i.

The weighting function (kernal) used in GWR can take different forms. The simplest approach applies a weight of 1 to all data within a set distance (d) of i and 0 to data outside of this area (Figure 3.2). This function creates artificial boundaries and could create artefacts in the patterns of parameters estimated. An alternative, illustrated in Figures 3.2 and 3.3, is to use a continuous weighting function such as a bi-square function that approximates a Gaussian weighting within the bandwidth and zero beyond this (Equation 3.8):

$$w_{ij} = [1 - (d_{ij}/\alpha)^2]^2 \text{ if } d_{ij} < \alpha$$

$$= 0 \text{ otherwise} \tag{3.8}$$

where the weight applied (w_{ij}) is a function of distance (d) between the central data point at i, a different data point at j and a parameter referred to as the bandwidth (α). The bandwidth defines the total distance of the subset of data (*Fotheringham et al.*, 2002) (Figure 3.3). The bandwidth is a global parameter and can be defined either as a fixed distance (fixed kernel) or by a number of nearest neighbours (adaptive kernel). The latter definition allows the weighting function to respond to changes in
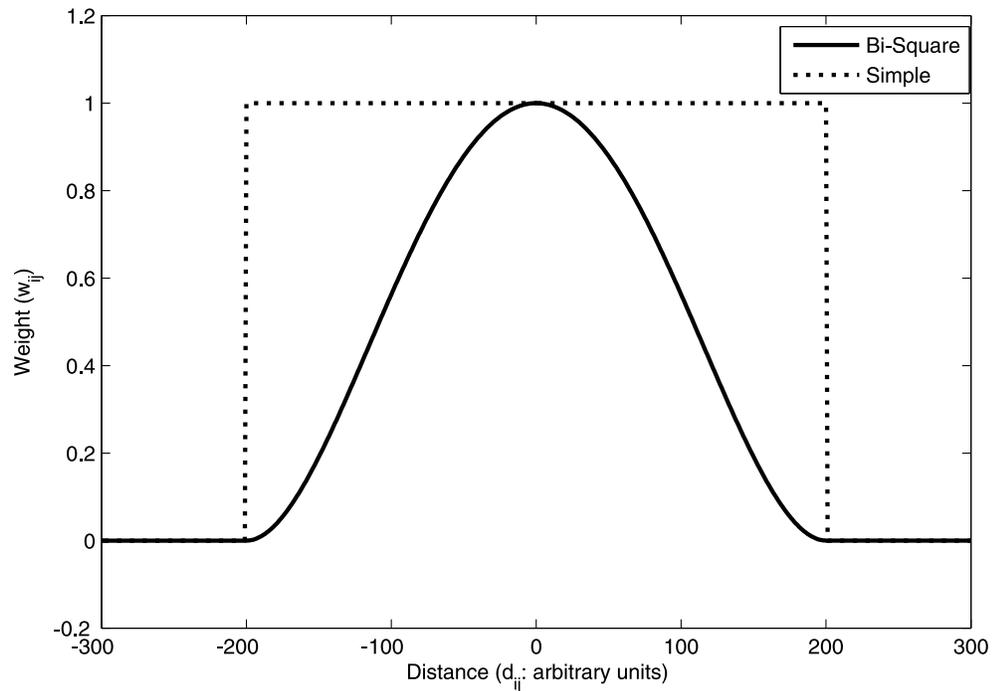
Figure 3.2: Example of the bi-square weighting function used in GWR from Equation 3.8. The function is shown using arbitrary distance units from a central data point and a arbitrary bandwidth (α)=200. As a comparison a simple weighting function is shown where all data are weighted equally within the bandwidth. Note that the dotted line at zero is covered by the full line.

sampling density where a fixed kernel may be unsuitable to apply. In both definitions the weighting function defines a 'bump of influence' around data at point i (Figure 3.3).

The size of the bandwidth is a critical factor in GWR as it defines the area of influence of each regression. A larger bandwidth will lose information about any spatial variability in coefficients if present and bias the results towards the global regression. If regression coefficients are considered to be a continuous field in space, then each data point in the subset of data will have a unique coefficient value, but defining a subset of data around each point forces the coefficient to a common (essentially an average) value for that subset. Therefore, coefficients can never be completely unbiased because there is always a level of spatial averaging. To minimise the bias of coefficients, a small subset of data close to i is preferable although this increases the variance and standard error of the regression model. There is, therefore, a trade-off between the increased variance associated with smaller subsets of data and the bias towards the global coefficients
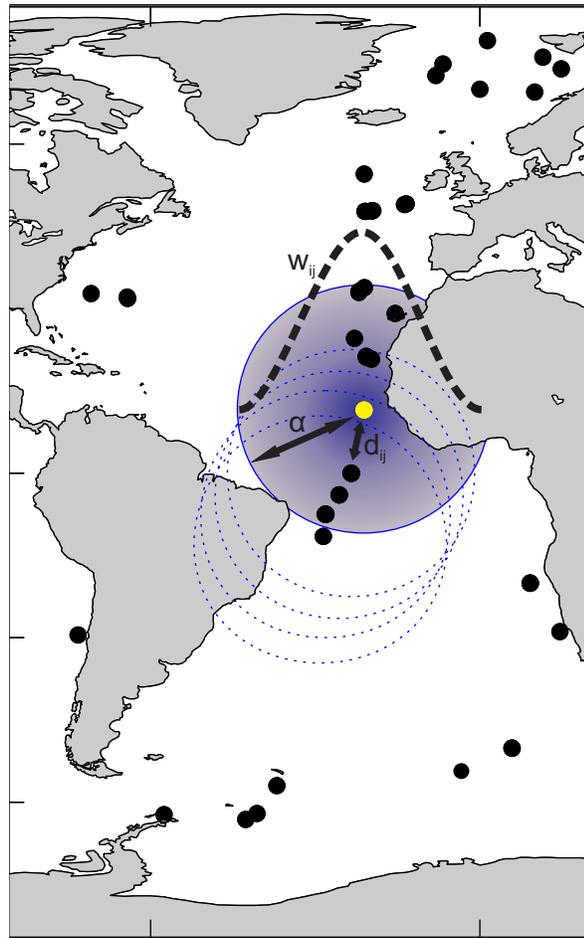
Figure 3.3: Schematic of the Geographically Weighted Regression process. The yellow data point indicates the central location of the analysis. The blue shaded circle indicates a bandwidth and the parameters that define it defined in equation 3.8. The dashed blue circles indicate previous iterations. Note that the schematic resembles the case where the bandwidth is defined as a fixed distance for ease of illustration. When applied to the sediment trap dataset a bandwidth defined by a number of nearest neighbours is used which is illustrated in Figure 3.8.

at larger subsets. To address this issue, the bandwidth can be calibrated using the cross validation score (CV) or the corrected Akaike Information Criterion (AICc: see *Akaike*, 1974). These quantities express the overall performance of regression models and can take into account the bias-variance trade-off, providing an estimate of the best bandwidth to use (*Fotheringham et al.*, 1998). For example, the AICc for regression models is a function of the residual sum of squares (AICc penalises larger residual sum of squares, *i.e.,* a worse fit to observations) and the number of data points (AICc penalises smaller datasets) (*Akaike*, 1974). The AICc therefore provides a quantitative

measure of the bias-variance trade-off.

A number of statistical tests are included with software used to calculate GWR to allow the assessment of GWR models against global regression models (for full details see *Fotheringham et al.*, 2002). These include Analysis of Variance (ANOVA), testing the null hypothesis that the GWR model represents no improvement on the global model, and a Monte Carlo test to assess the significance of the spatial variability of GWR coefficients. Under the null hypothesis, any random permutation of the dataset is equally likely to occur. The details of this output can be found in Appendix A.

3.2.2.3   *Application of GWR to Sediment Trap Data*

Geographically Weighted Regression provides a rigorous approach to assessing the spatial variability of carrying coefficients. Although there has been extensive work on extending the predictive use of GWR (*Harris et al.*, 2011; *Kumar and Lal*, 2011), GWR is applied here as an exploratory tool. Other related spatial statistics, such as Kriging, could similarly be applied to the sediment trap data. The advantage of the GWR approach described here is that the original multiple linear regression model used by *Klaas and Archer* (2002) is directly comparable to the results described in this chapter.

GWR is applied to the updated sediment trap dataset using software (GWR 3.0) kindly provided by M. Charlton of the National Centre for Geocomputation, National University of Ireland Maynooth. An adaptive kernel (defining a subset of data by number of nearest neighbours) was chosen to group data because a fixed kernel (defining a subset of data strictly by distance) failed to produce results due to the sparse sampling density in space. For example a fixed kernel large enough to sample reliably in the Southern Ocean would be too large to reveal spatial relationships in other regions. Conversely, a fixed kernel able to reveal spatial variability in some regions would sample too little data in the sparsely sampled Southern Ocean. The bandwidth was chosen by finding the minimum AICc score using an optimisation routine in the GWR 3.0 software
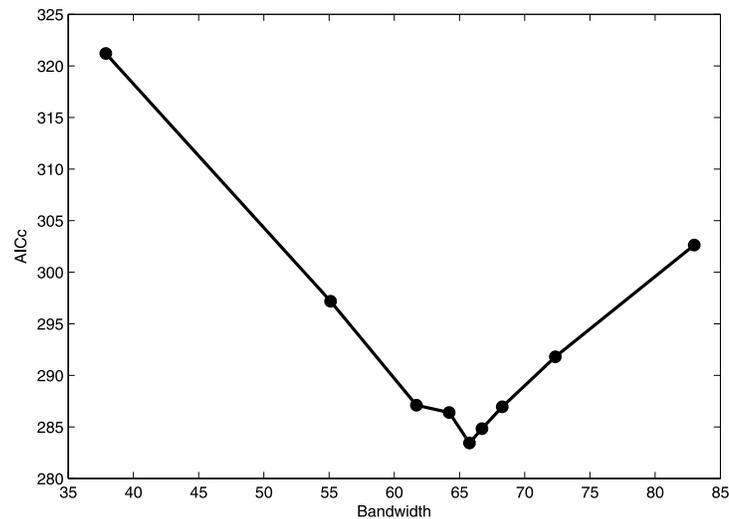
Figure 3.4: The AICc minimisation technique used to find the optimal bandwidth when applied to the sediment trap dataset. The optimal bandwidth is 66 nearest neighbours as indicated by the minimum in AICc.

(Figure 3.4). When applied to the sediment trap dataset, a bandwidth of 66 nearest neighbours is found. Data were weighted using bi-square function (Equation 3.8). A Monte Carlo significance test also provided with the GWR 3.0 software was used to determine whether the regression coefficients were spatially variable (*Hope*, 1968).

3.3 Results

3.3.1 Global Regression

The sediment trap dataset shows similar global relationships to those obtained previously with smaller datasets (Figure 3.5 *cf.* 1.7). The correlation between POC and $CaCO_3$ is strongest (r=0.60), whilst the correlation between POC and opal is weaker (r=0.35), with lithogenic fluxes being intermediate (r=0.45). Visually, the association of POC with both opal and lithogenic material (Figure 3.5) suggests the presence of two separate distributions, one of high POC flux and low mineral flux and the second of low POC flux with high mineral flux (this was originally noted for opal by *Klaas and Archer*, 2002). The scatter plot for $CaCO_3$ also displays more variability

than previously recognized (Figure 1.7). All scatter plots display regional differences when separated into ocean basins as previously noted by *Ragueneau et al.* (2006) and *De La Rocha and Passow* (2007). The global POC:mineral ratio is 0.052, close to the ratio observed by *Armstrong et al.* (2002).

Multiple linear regression is used on the global dataset to express the flux of POC as a summed linear function of mineral fluxes, and, as observed previously, suggests a dominant role for $CaCO_3$ (Table 3.2). The resulting regression model is significant at $p<0.001$ and predicts 66% of the variability in POC fluxes ($R^2=0.66$). The carrying coefficient for $CaCO_3$ is close to, although very slightly higher than, previous estimates (Table 3.1). The coefficient and standard errors for opal are also similar in magnitude to previous estimates whilst the lithogenic coefficient is much smaller than that found by *Klaas and Archer* (2002) and *Francois et al.* (2002). Both the estimate here and that of *Ragueneau et al.* (2006) are derived from significantly larger datasets (n=156-189 *cf.* n=62-78) suggesting that this value should be more globally representative.

The R^2 value is relatively low in comparison to the global studies in Table 3.1. This is due to the inclusion of the intercept term in equation (3.5). Repeating the analysis without the intercept term increases R^2 to 0.89, a value consistent with previous studies (Table 3.1). The large difference between these values is due to the calculation of R^2 from the total sum of squares which is uncorrected without an intercept, leading to inflated values of R^2 (*Montgomery et al.*, 2006). To be consistent and directly comparable with geographically weighted regression analysis (GWR 3.0 does not allow for a model without an intercept) an intercept is included, justified by the fact the residual mean squares (MSE) are very similar (MSE with intercept=0.484, MSE without intercept=0.481) and that the derived carrying coefficients only differ by a factor of 0.0001. Note, therefore, that R^2 values in this study are not directly comparable to previous studies.
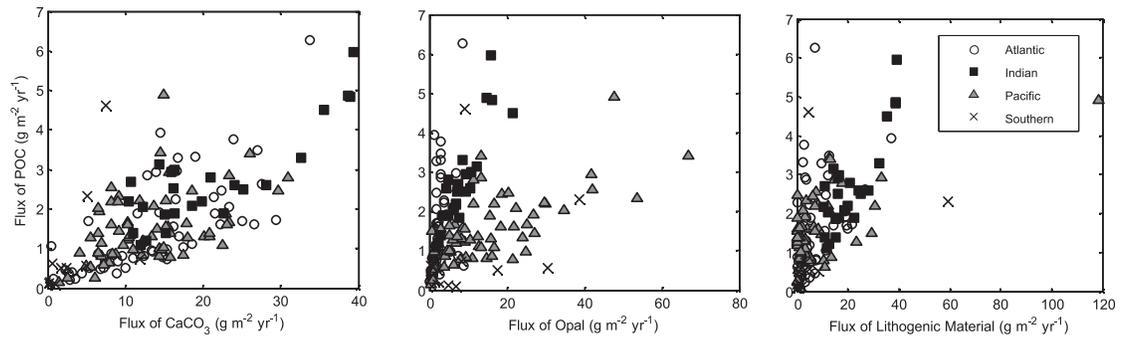
Figure 3.5: Global annual mass fluxes of particulate organic carbon versus global annual mass fluxes of (left) $CaCO_3$, (middle) opal and (right) lithogenic material as measured by sediment traps >1500m depth. Flux data from different ocean basins are indicated by symbols to highlight regional differences (adapted from *Ragueneau et al.*, 2006).

Table 3.2: Carrying coefficients calculated using multiple linear regression with mass flux data for different depth ranges and spatial scales. All coefficients are significant at p<0.001 except a where p<0.05. 95% confidence intervals are given as 2x standard error. Note that the model for the Southern Ocean is not significant.

Global Data	$CaCO_3$	Opal	Lithogenic	R^2
>1500m (n=156)	0.089	0.023	0.027	0.66
	0.076-0.102	0.012-0.034	0.017-0.037	
Regional Data (>1500m)				
Atlantic (n=54)	0.083	0.152^a	0.027^{NS}	0.58
	0.047-0.118	0.028-0.276	-0.007-0.060	
Indian (n=25)	0.083	0.058^a	0.058	0.94
	0.058-0.108	0.003-0.113	0.034-0.083	
Pacific (n=63)	0.056	0.033	0.022	0.80
	0.041-0.071	0.025-0.041	0.015-0.028	
Southern Ocean (n=12)	0.183^{NS}	-0.022^{NS}	0.034^{NS}	0.37^{NS}
	-0.065-0.431	-0.139-0.095	-0.050-0.117	

3.3.1.1   Evidence of Spatial Variability

A regional breakdown of the data to an ocean basin scale is also shown in Table 3.2 and is comparable to that by *Ragueneau et al.* (2006) as summarized in Table 3.1. The carrying coefficient for opal in the Atlantic displays variability compared to other basins and is quite distinct from the global value (0.152 *cf.* 0.023). In the Pacific and Indian basins, the spatial coefficients are more consistent with the global values. The Indian basin values differ to those found in Table 3.1 the reason for which is unclear. The relevant dataset in Table 3.1 is much smaller (n=16 *cf.* n=25) and it may be that these coefficients are more influenced by outliers in the regression. In particular, it is also worth considering that the Pacific basin is the largest in size, covering 19 biogeochemical provinces and potentially includes multiple sources of variability unlike the smaller Indian basin comprising 8 biogeochemical provinces. Therefore, it is uncertain whether the similarity of the Pacific coefficients to global values may be a product of averaging the potentially large spatial variability in fluxes or is actually representative of the values in this area. Table 3.2 also includes analysis of data from the Southern Ocean but the resulting regression model is not significant. Visually there are differences within the Southern Ocean in Figure 3.5 particularly for opal which may suggest variability between different sectors. It is worth noting the majority of the sediment trap data is located in the Atlantic sector of the Southern Ocean (Figure 3.1) although any bias this may have is difficult to assess given the very small sample size (n=12). These results highlight the difficulties in conducting regional regression even at the scale of ocean basins. Making comparisons between areas in this way is problematic because of varying sample sizes and a lack of consistent statistical methodology.

An alternative method of assessing whether regional variability exists in the global dataset is to map the residuals of the global regression model (Figure 3.6a). An assumption of regression is that the residuals should have a random distribution around zero. Extending this logic, if coefficients are truly global they should also exhibit resid-

uals that are randomly distributed in space (*Fotheringham et al.*, 2002). However, note here that negative residuals appear to cluster in the low-latitude Atlantic and Pacific as well as the western sub-arctic Pacific whereas positive residuals cluster in the Arabian Gulf and Indian Ocean. This supports the contention from previous studies that there is potential spatial non-stationarity in the coefficients which are not truly global (*Passow and De la Rocha*, 2006; *Ragueneau et al.*, 2006).

3.3.2 Geographically Weighted Regression

Geographically weighted regression analysis was applied to the global sediment trap dataset at depths >1500m. The AICc calibration was used to determine an optimal bandwidth of 66 nearest neighbours (Figure 3.4).

3.3.2.1   Assessing Geographically Weighted Regression

A reduction in the AIC score from the global model score of 341.6 and an increase in R^2 from 0.66 to 0.82 respectively suggest that GWR is an improvement on the global model (Table 3.3). The results of the ANOVA statistics also show that the GWR model is a significant improvement on the global model whilst the results of the Monte Carlo test suggest there is significant spatial variability in the regression coefficients. A visual comparison of the residuals from the GWR (Figure 3.6b) against the global regression residuals (Figure 3.6a) indicates greater heterogeneity in areas previously characterised by clustered residuals, such as the equatorial Pacific, Atlantic and the Indian Ocean. These metrics suggest that the use of a spatially informed regression technique is justified here.

GWR calculates regression coefficients and other statistics at each data point, allowing them to be mapped (Figure 3.7). The results show distinct regional groupings of the coefficients with minimal variability within these groups. An exception to this are the coefficients for opal in the North Atlantic, which show a range of values in a relatively
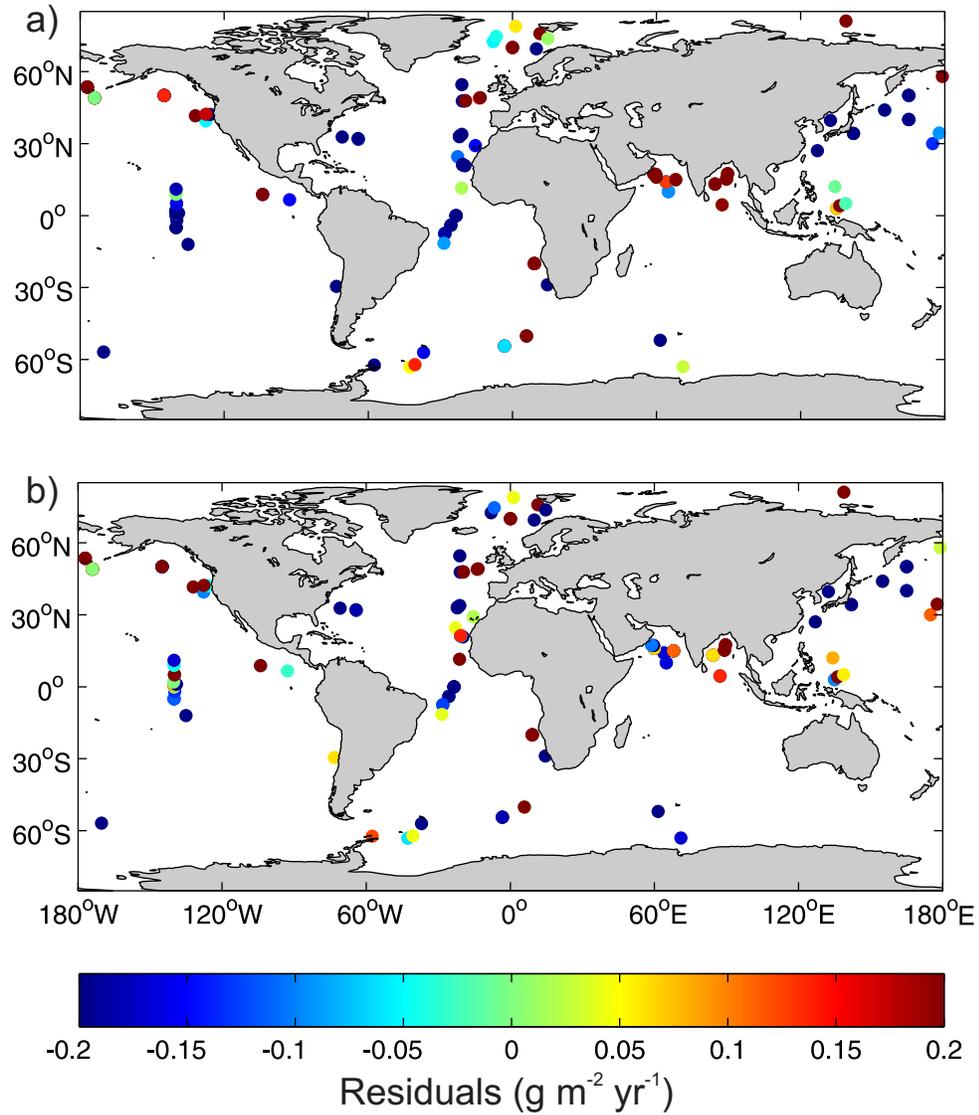
Figure 3.6: The spatial distribution of residuals from (a) global multiple linear regression and (b) geographically weighted regression analysis. Fluxes of POC are predicted using mass fluxes of $CaCO_3$, opal and lithogenic material as in *Klaas and Archer* (2002). The models predict ∼66% and ∼82%, respectively, of the variability in observed POC fluxes.

Table 3.3: Summary statistics of the GWR models for sediment trap data at >1500m. A bandwidth of 66 is the optimal value chosen from the AICc calibration. Additional bandwidths of 20 and 156 are provided to explore the sensitivity of the outcomes to bandwidth. AIC is an comparative quantitative score of the regression models where lower values indicate better models. ANOVA statistics test the null hypothesis that the GWR model is no improvement on the global model. 'Global' carrying coefficients are the median of the 156 coefficients. Monte Carlo test for spatial variability rejects the null hypothesis at *p<0.05, **p<0.01,*** p<0.001

Bandwidth	AIC	R^2	ANOVA Sig.	$CaCO_3$	Opal	Lithogenic
66 (optimal)	269.82	0.82	<0.001	0.066***	0.033***	0.022***
20	318.04	0.90	<0.001	0.056	0.040***	0.026***
156	313.41	0.71	<0.001	0.089***	0.025	0.026

small area with no identifiable spatial trend. In this region, the bandwidth of 66 nearest neighbours that defines the subsets of data is relatively large in comparison to the area (see Figure 3.8a), and is therefore influenced by significantly different values in the Arctic and equatorial Atlantic Oceans. This highlights a particular issue of using GWR with this dataset: the analysis is limited by the relatively small number of data points compared to the area sampled, such that areas of low sampling density may produce spurious results. Data points at the edge of ocean basins or in sparsely sampled areas, such as the Southern Ocean, will also tend to include data from other basins. The effect of sparsely sampled areas can be seen in the local R^2 values (the fit to each subset of data). In the Southern Ocean, the R^2 is the lowest of the dataset reflecting that data from potentially different regions is included (Figure 3.8b). The inclusion of data from other ocean basins is an important caveat for this analysis and limits the following section to the dicussion of large-scale spatial patterns in coefficients. This caveat is hereafter referred to as inter-basin influence.

To explore the sensitivity of coefficients to the bandwidth and inter-basin influence, the bandwidth was manually changed to 20 and 156 nearest neighbours in comparison to the calibrated 66 neighbours (Table 3.3) and the results from the most variable mineral, $CaCO_3$, were plotted (Figure 3.9b & 3.9c). A bandwidth of 156 is used to assess if the GWR technique can recover MLRA global coefficient values. Although
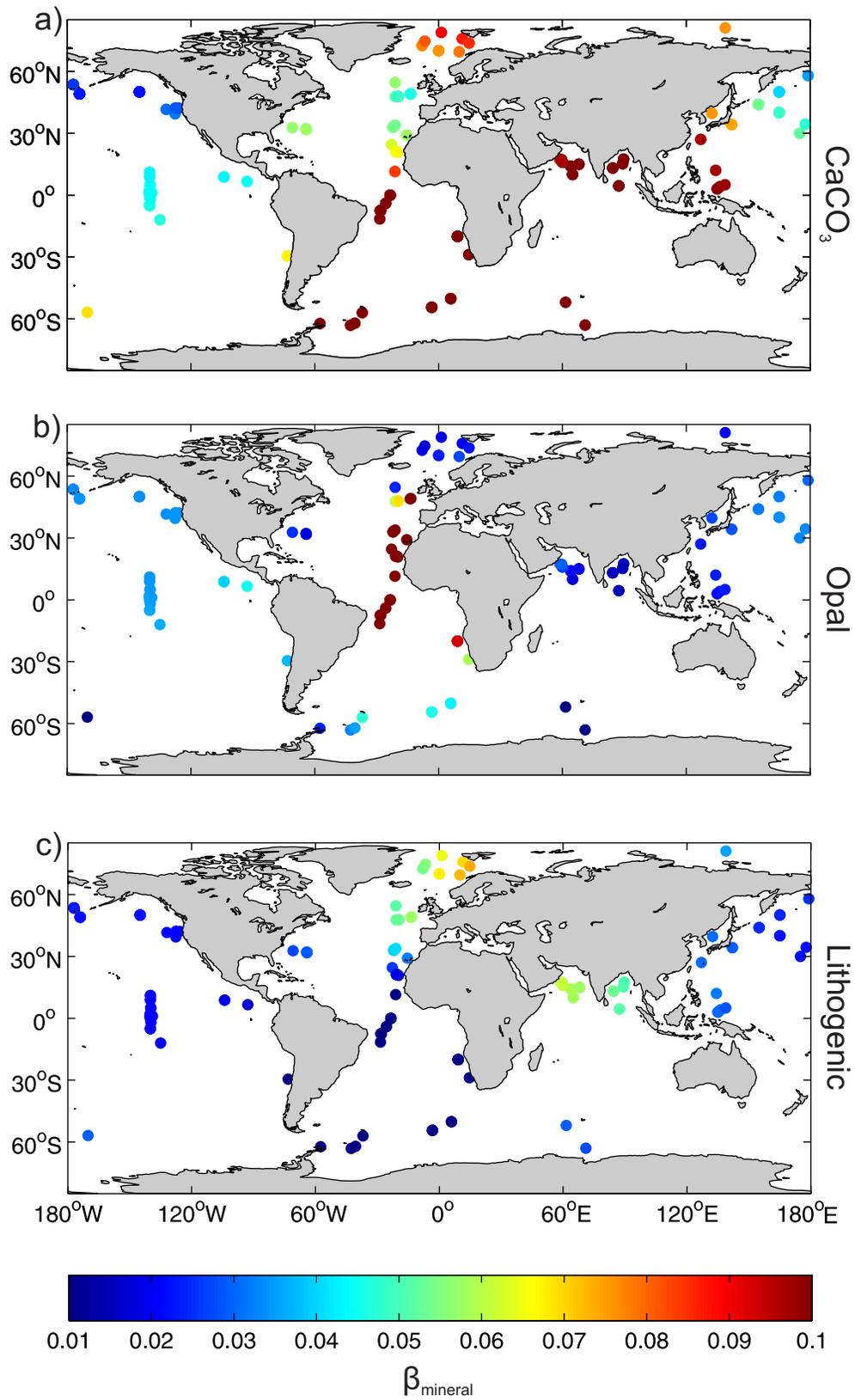
this will include all data points in the dataset they will still be weighted by distance. Figure 3.9 shows that most of the coefficients converge towards a global value around 0.09. The equatorial Pacific is the only area not to conform to the global value. This is likely to be a result of a combination of densely sampled data being weighted heavily but in relative isolation to other data points. The higher AIC score suggests this model has less ability to predict POC than the calibrated optimal model, which is also suggested by the lower R^2 value (0.71) although it is still higher than the global mean MLRA model (0.66) (Table 3.1). With a much smaller bandwidth of 20, the R^2 value increases but the higher AIC score indicates other aspects of the model are worse, such as increased variance (Table 3.3). One might expect coefficients to vary significantly at this bandwidth because the coefficients represent local values but are also subject to influence from outliers. However, the large-scale general spatial patterns (Figure 3.9b) are comparable to those in Figure 3.7a, suggesting these are not an artifact of bandwidth size. A notable exception to this is the lower coefficients in the equatorial Atlantic (Figure 3.9b *cf.* Figure 3.7a). This suggests that the subset of data defined by a bandwidth of 66 nearest neighbours may be influenced by significantly higher values in the Southern Ocean but overall our ability to capture the global mean MLRA coefficient values when relaxing the spatial bandwidth gives us increased confidence in this method.

The basin-scale coefficients in Table 3.2 appear to corroborate the GWR coefficients. Furthermore, to test the validity of large scale patterns of coefficients produced by GWR, subsets of data were manually selected by region and compared them to the

---

Figure 3.7 *(preceding page)*: The spatial distribution of coefficients calculated using geographically weighted regression analysis for (a) $CaCO_3$, (b) opal, and (c) lithogenic material. The model is the same multiple linear regression model applied to the global data in Table 3.2 and Figure 3.6b. The GWR analysis uses a bandwidth of 66 nearest neighbours, defined by an AICc minimization calibration procedure. A bi-square weighting scheme was used. The GWR model predicts ~82% of the variability in POC fluxes and is an improvement on the global model. The corresponding global coefficient values are 0.089, 0.023, and 0.027 for (a), (b), and (c) respectively.
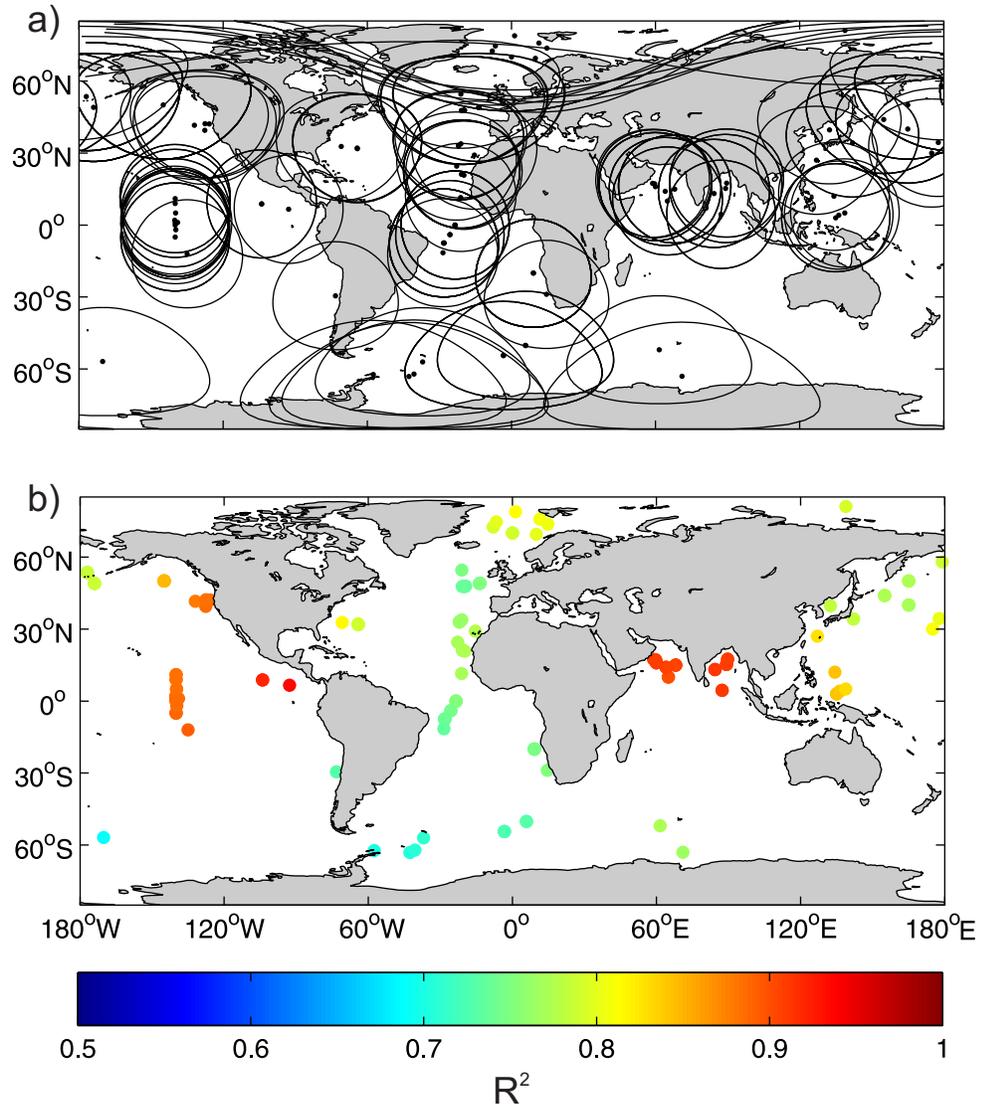
Figure 3.8: a) Schematic depicting the bandwidth of geographically weighted regression. Circular line indicated the distance at which a weight of 0.75 occurs. Data is weighted >0.75 within the area and <0.75 up to the distance of the last nearest neighbour. A bandwidth of 66 nearest neighbours is used. (b) Local R^2 values from the geographically weighted regression in Figure 3.7 using a bandwidth of 66 nearest neighbours and bi-square weighting scheme.
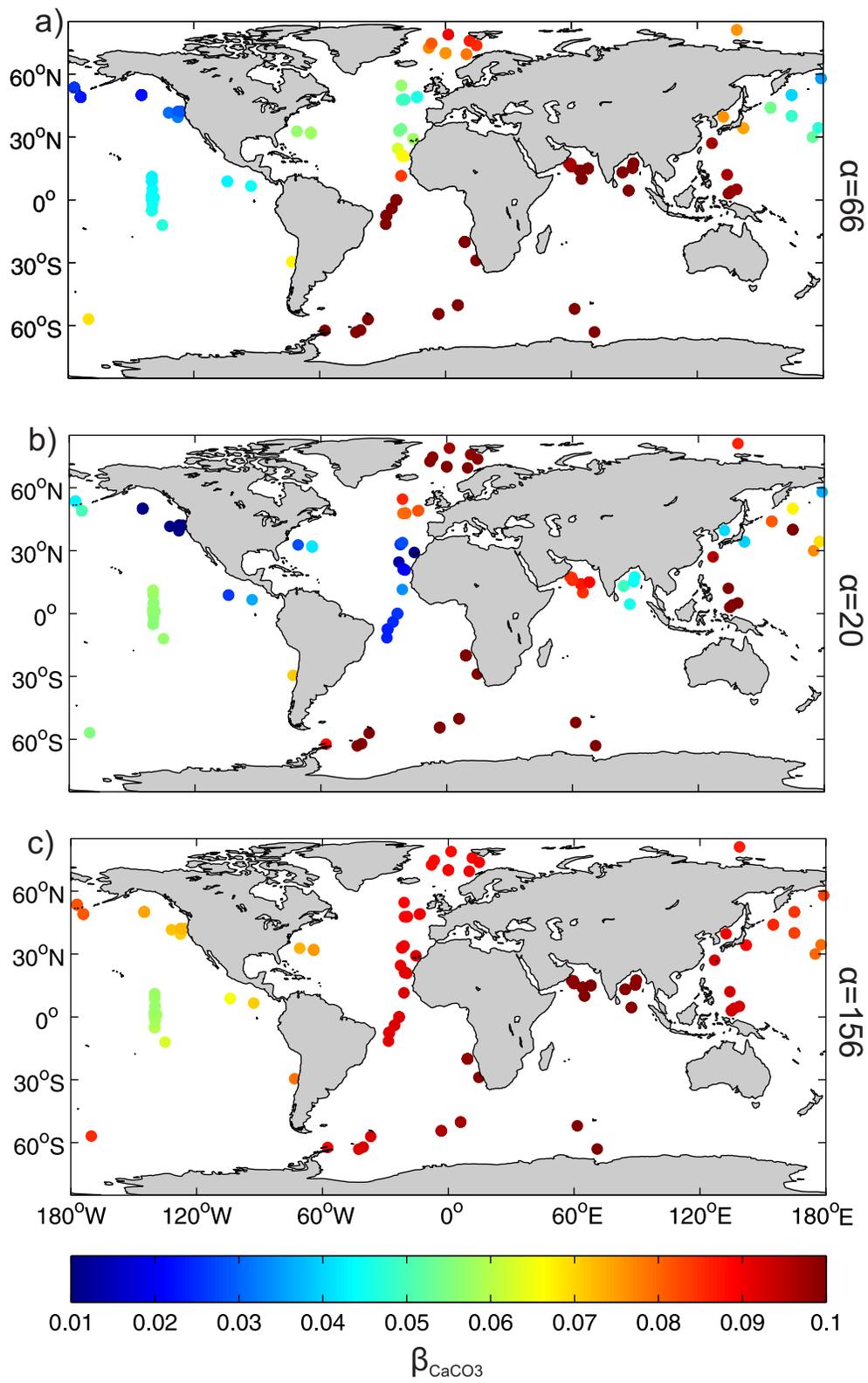
mean GWR coefficients for the corresponding area (Figure 3.10). Overall there is general agreement between the coefficients from both methods, supporting the results of the GWR analysis. This method of validation is difficult because some areas have a small number of data points, which can lead to statistically unreliable outcomes and is one reason why GWR is a preferable technique. Overall, the coefficients show general agreement with the GWR coefficients, suggesting that the GWR technique is producing reliable results even when considering the potential for inter-basin influence.

3.3.2.2   *Spatial Patterns in Regression Coefficients*

The geographically weighted regression analysis defines clear regional patterns in the coefficients for $CaCO_3$, opal and lithogenic particles (Figure 3.7). The Atlantic displays some of the more unexpected results, with coefficients differing appreciably from global values. Coefficients for both $CaCO_3$ and opal display a decreasing trend with increasing latitude between 30 and 60 degrees North in the North Atlantic. Opal is quantitatively more important than $CaCO_3$ and lithogenics in the low latitude Atlantic, corroborating the basin-scale approach in Tables 3.1 & 3.2. In the Arctic Ocean however, $CaCO_3$ and lithogenics are quantitatively more important. This is not revealed by the basin-scale results.

Overall in the Pacific, there is much less difference between coefficient values, with no mineral showing overall importance except for the higher coefficients observed for $CaCO_3$ in the western and equatorial Pacific. These could be a result of inter-basin influence from the Indian Ocean although similar behavior is not observed in the lithogenic coefficients, suggesting that it is probably not an artefact. The majority of the $CaCO_3$

Figure 3.9 *(preceding page)*: Spatial distribution for $CaCO_3$:POC coefficients calculated using geographically weighted regression (a) from Figure 3.7 but with manually altered bandwidth values of (b) 20 and (c) 156 nearest neighbours. Only coefficients for $CaCO_3$ are displayed, as it displays the most variability of the three minerals considered. For reference, the global $CaCO_3$ coefficient is 0.089.
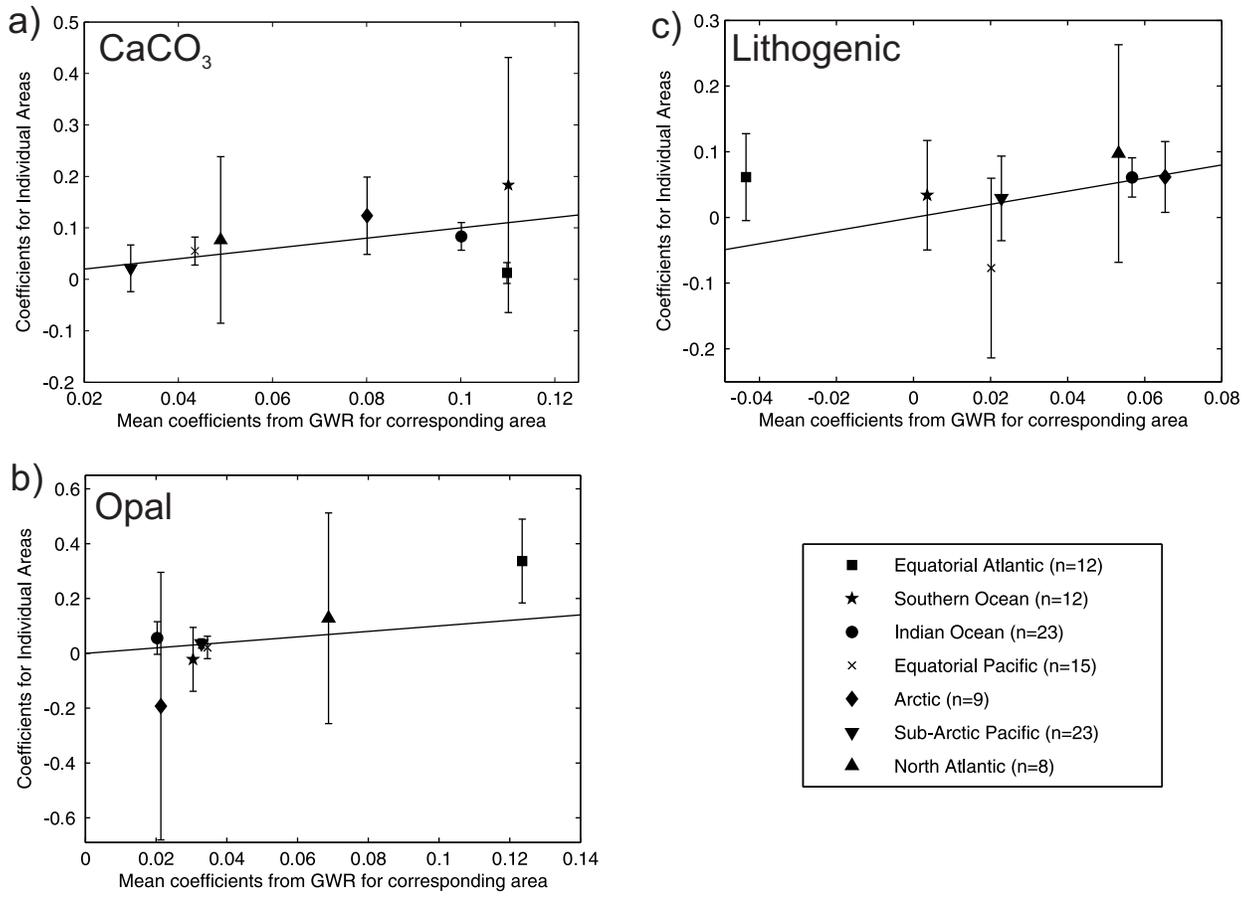
Figure 3.10: Comparison of regression coefficients calculated over manually selected areas and mean coefficient values calculated using geographically weighted regression for the corresponding area for (a) CaCO$_3$, (b) opal, and (c) lithogenic material. Error bars correspond to 2 x standard error; 1:1 ratio lines are shown.

coefficients in the Pacific are much lower than the global value of 0.089 (as also highlighted in the basin-scale analysis; Table 3.2).

The Indian Ocean is one of the only regions that display similar coefficients to those from the global analysis, with a quantitative importance identified between both CaCO$_3$ and lithogenics, and POC. Variability between the Arabian Gulf and Bay of Bengal is small and is likely a result of subtle changes in the subsets of data used. Finally, the Southern Ocean also displays unexpected results, with quantitative importance between CaCO$_3$ and POC only. Inter-basin influence may be a large problem for the Southern Ocean, given the small number of samples and the relative distances between them (see Figure 3.8a for indication of bandwidth sizes). This is reflected in both the insignificant

regression result in Table 3.2 and the relatively low R^2 values generated by the GWR analysis (Figure 3.8b).

Overall the global coefficient for $CaCO_3$ of 0.089 appears to be dominated by the Southern Ocean, the Indian Ocean, western equatorial Pacific and the Arctic whilst regionally displaying much lower values (<0.04-0.05) in the Atlantic and the remaining Pacific. In contrast, the global values of 0.023 for opal and 0.027 for lithogenics appear more consistent over large regions, with exceptions in the Atlantic for opal and the Artic and Indian Oceans for lithogenics.

3.3.2.3   *Predictions of POC fluxes*

The carrying coefficients for each mineral in the regression analyses predict the flux of POC to depth as a function of ballasting fluxes. Figure 3.11 compares the POC flux predicted by both the MLRA and GWR models as a fraction of the total POC flux predicted, as suggested by *Le Moigne et al.* (2012). Globally, the GWR model predicts less POC associated with $CaCO_3$ than the MLRA model. This is predominantly a result of differences in the Equatorial and North Pacific as well as the North Atlantic (Figure 3.11a *cf.* Figure 3.11b). Since the MLRA coefficient for $CaCO_3$ is a single global coefficient, it follows that the predicted POC flux mirrors the $CaCO_3$ fluxes, *i.e.,* ballasting. However, GWR predictions suggest that there is no clear relationship between POC and $CaCO_3$. There are less differences between the two regression models for opal and lithogenic material, with greater POC fluxes in the North Pacific and the Bay of Bengal respectively (Figure 3.11c & 3.11e *cf.* Figure 3.11d & 3.11f).

3.3.3   Sampling Artefacts: Synthetic Flux Comparison

Sediment trap flux measurements can potentially be biased by a number of factors. Attempts to reduce biases due to undersampling and zooplankton have been taken by not considering sediment traps <1500m. An additional source of uncertainty in the use
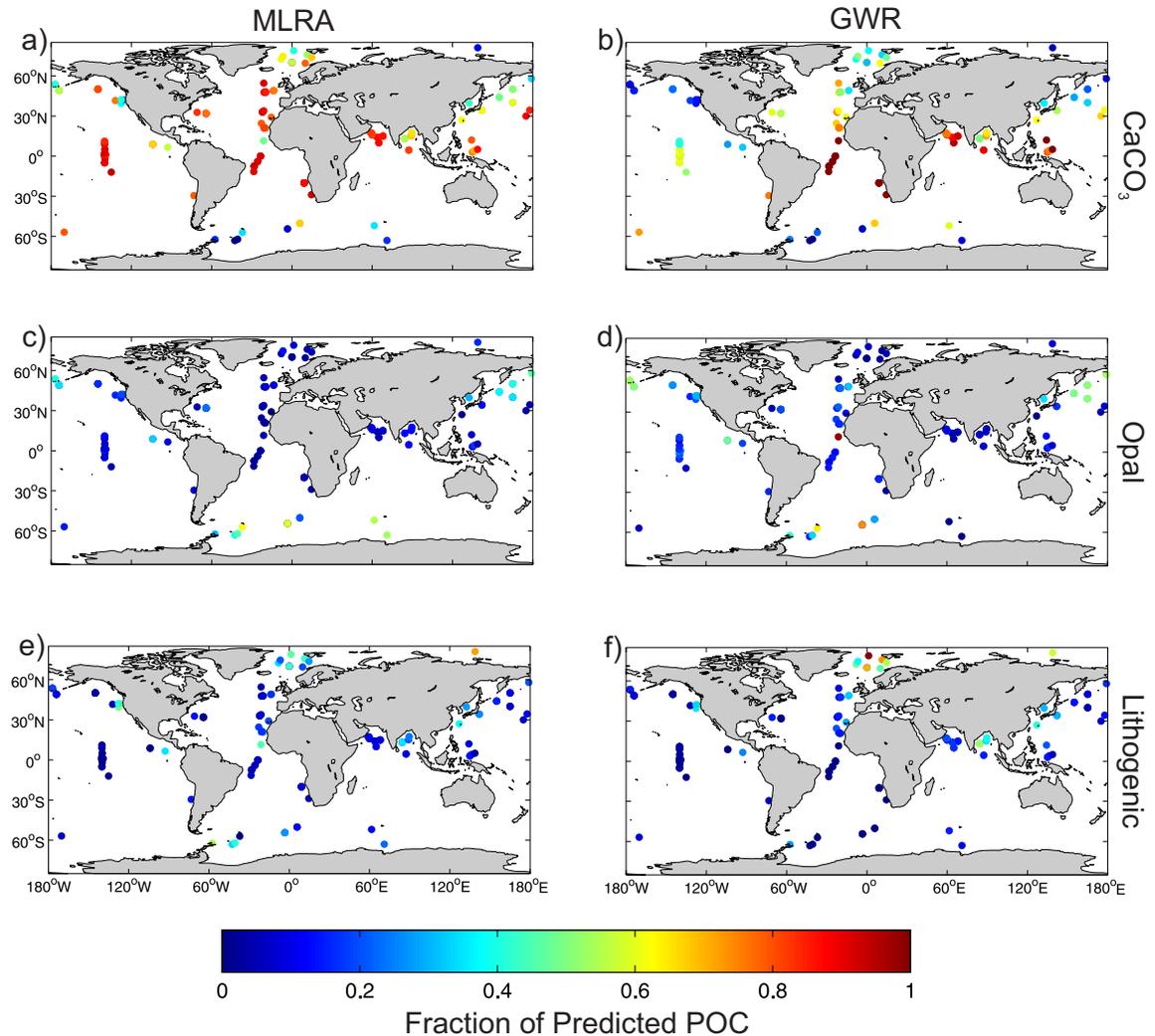
Figure 3.11: MLRA and GWR predictions of POC fluxes from each mineral. Values are expressed as the contribution from each mineral ($\beta_{CaCO_3} \cdot F_{CaCO_3}$) as a proportion of the total predicted POC flux, following *Le Moigne et al.* (2012). The GWR model is the corresponding model in Figure 3.7.

of deep sediment traps is the dissolution of samples once sampled due to the effects of the ambient water chemistry, *e.g.,* low carbonate ion or dissolved silicic acid concentrations (*Buesseler et al.*, 2007). This effect will occur in particular regions such as the deep North Pacific, where carbonate ion concentrations are low due to the accumulation of dissolved CO_2 (*Sarmiento and Gruber*, 2006). In contrast, the ocean is under-saturated everywhere with respect to biogenic opal. The $CaCO_3$ coefficients in the North Pacific are the smallest in magnitude and form a distinct regional grouping (Figure 3.7a) suggesting that a dissolution bias should be considered. To address this, GWR is applied

to a set of synthetic fluxes that are estimated from an empirical model that explicitly excludes the effects of dissolution. The use of a synthetic dataset provides a qualitative test that the first-order variability in coefficients calculated from the observed flux dataset are attributable to changes in the sinking particulate flux and not locally-occurring dissolution.

Asynthetic dataset is derived from the studies of *Dunne et al.* (2005), *Dunne et al.* (2007) and *Dunne et al.* (2012). These studies present a synthesis of empirical and mechanistic models of productivity, export, and sinking particulate fluxes that are applied to satellite chlorophyll measurements to estimate the oceanic cycles of organic carbon, silicon and $CaCO_3$. Other methods were considered, such as using ocean biogeochemical models to estimate fluxes, but the use of ratios to predict $CaCO_3$ from POC precluded a lot of spatial variability. The use of more dynamic ecosystem models, *e.g., Yool et al.* (2013), may facilitate similar analysis. Estimates of the POC exported from the surface are derived from an export production algorithm based on sea surface temperature (*Dunne et al.*, 2005) applied to estimates of primary production estimated from 1° x 1° satellite measurements of chlorophyll, *e.g., Behrenfeld and Falkowski* (1997). Two algorithms are then used to estimate POC fluxes at depth (2000m); a fixed Martin Curve where the exponent $b = $ -0.858 (see Section 1.2.1), and spatially variable exponents from *Henson et al.* (2012). Export fluxes of opal are taken from *Dunne et al.* (2007), which are based on the method of inferring export fluxes from vertical surface and subsurface nutrient gradients (*Sarmiento et al.*, 2004a). This method predicts silica export:organic carbon export ratios in 10 different regions. Opal fluxes at depth are estimated using an exponential decay curve with a length scale of 2000m (*Dunne et al.*, 2007). Fluxes of $CaCO_3$ are taken as the estimated bottom fluxes of *Dunne et al.* (2012), again derived from satellite chlorophyll using a variety of algorithms. The POC and opal export fields are taken directly from the publications and estimated at depth separately whilst the bottom fluxes of $CaCO_3$ are taken directly from *Dunne et al.* (2012). All fields were sampled from the 1° x 1° grid-box closest to the location of the sediment traps used for
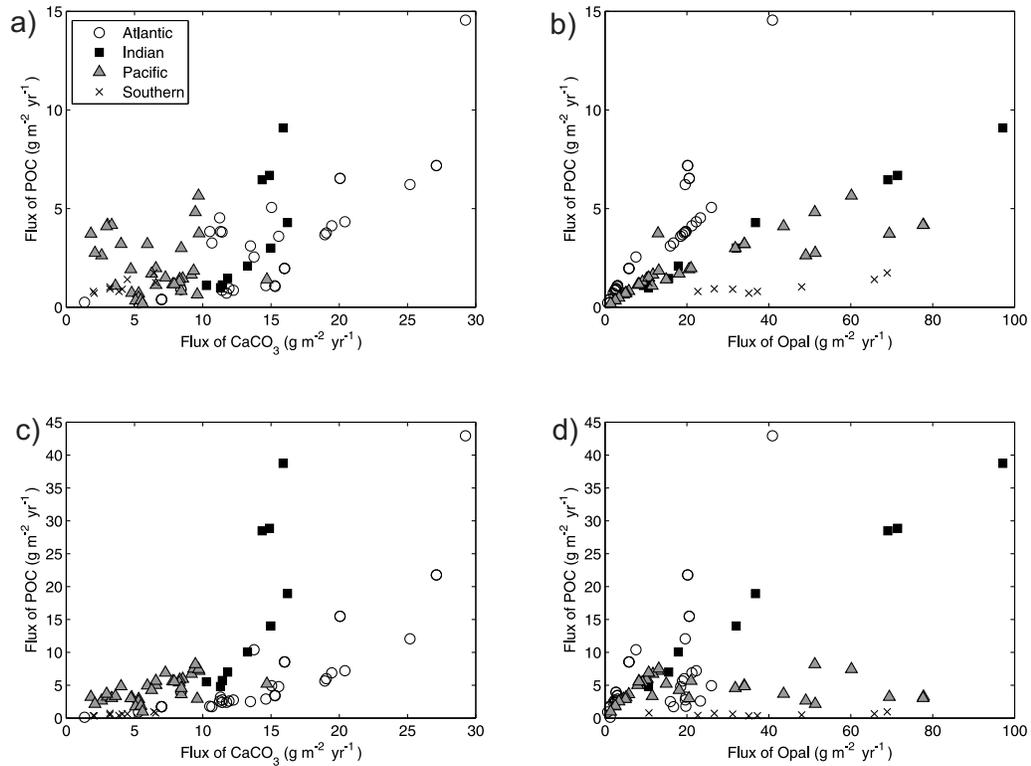
Figure 3.12: Estimated mass fluxes at depth (2000m) of particulate organic carbon versus CaCO$_3$ and opal based on *Dunne et al.* (2007, 2012). Panels (a) and (b) are calculated using a Martin Curve with a fixed global exponent. Panels (c) and (d) are calculated using the spatially variable exponents from *Henson et al.* (2012). This figure is directly comparable with Figure 3.1b.

the regression analyses (Figure 3.1a). A synthetic lithogenic dataset is not considered here as although estimates of dust deposition to the ocean surface are available, *e.g.*, *Mahowald et al.* (2005), the definition of lithogenic material in equation 3.4 as a residual flux, allows for additional sources such as riverine input that cannot be accounted for in a synthetic lithogenic dataset based on atmospheric dust deposition.

The synthetic datasets are shown in Figure 3.12. Prominent features of the sediment trap data (Figure 3.5) are also present in the synthetic dataset, such as a stronger global linear relationship between POC and CaCO$_3$ (Figure 3.1a and c) and the presence of the separate distributions for POC and opal (Figure 3.12b and d). The strong regional linear relationships in Figure 3.12b are an artefact of the algorithm used to estimate opal export by relating it to POC export by ratios in 10 regions. Overall the magnitudes of the fluxes are consistent with the sediment trap data, although some POC

fluxes calculated using variable Martin Curves are significantly higher in the Indian Ocean (Figures 3.12c and d). This is an artefact of the algorithm from *Henson et al.* (2012) which predicts a very large fraction of exported POC reaching the deep ocean ($\sim$45%) (see Figure 1.6) which is not supported by sediment trap observations that show smaller fractions (5-10%) (*Honjo et al.*, 2008). The global carrying coefficients for the synthetic dataset when using the constant Martin Curve (Figure 3.12a and Figure 3.12b) are 0.251 and 0.0671 for $CaCO_3$ and opal respectively (R^2=0.80) and when using the variable Martin Curves (Figure 3.12c and Figure 3.12c) are 0.913 and 0.159 (R^2=0.67). The differences in magnitude between these coefficients are those calculated from the sediment trap data likely reflect differences in flux magnitudes as noted above but both sets of coefficients indicate a strong quantitative relationship between POC and $CaCO_3$ as observed.

The same GWR analysis is applied as with the sediment trap data using an adaptive bandwidth of 66 nearest neighbours producing a corresponding set of coefficients (Figure 3.13). The general spatial patterns observed in the original analysis are reproduced relatively well using both synthetic datasets. In particular, the pattern of $CaCO_3$ coefficients in the North Pacific and Southern Ocean are consistent suggesting that dissolution within deep sediment traps before collection does not contribute to these patterns. Equally, both datasets broadly reproduce the relatively higher opal coefficients in the Atlantic. The dataset based on the globally constant Martin Curves captures more of the features seen in the observations such as the gradient of $CaCO_3$ coefficients in the Atlantic. Given first order estimates of flux magnitudes and distributions, the broad similarity of these results to the spatially variable coefficients in Figure 3.7 suggests that dissolution after sampling by sediment traps does not contribute to the patterns observed. Although, the synthetic datasets cannot be used a quantitative test of whether the carrying coefficients are a result of a ballasting mechanism, as this mechanism was not explicitly used to calculate the fluxes, it does suggest that the global and spatially varying carrying coefficients can be reproduced in the absence of a
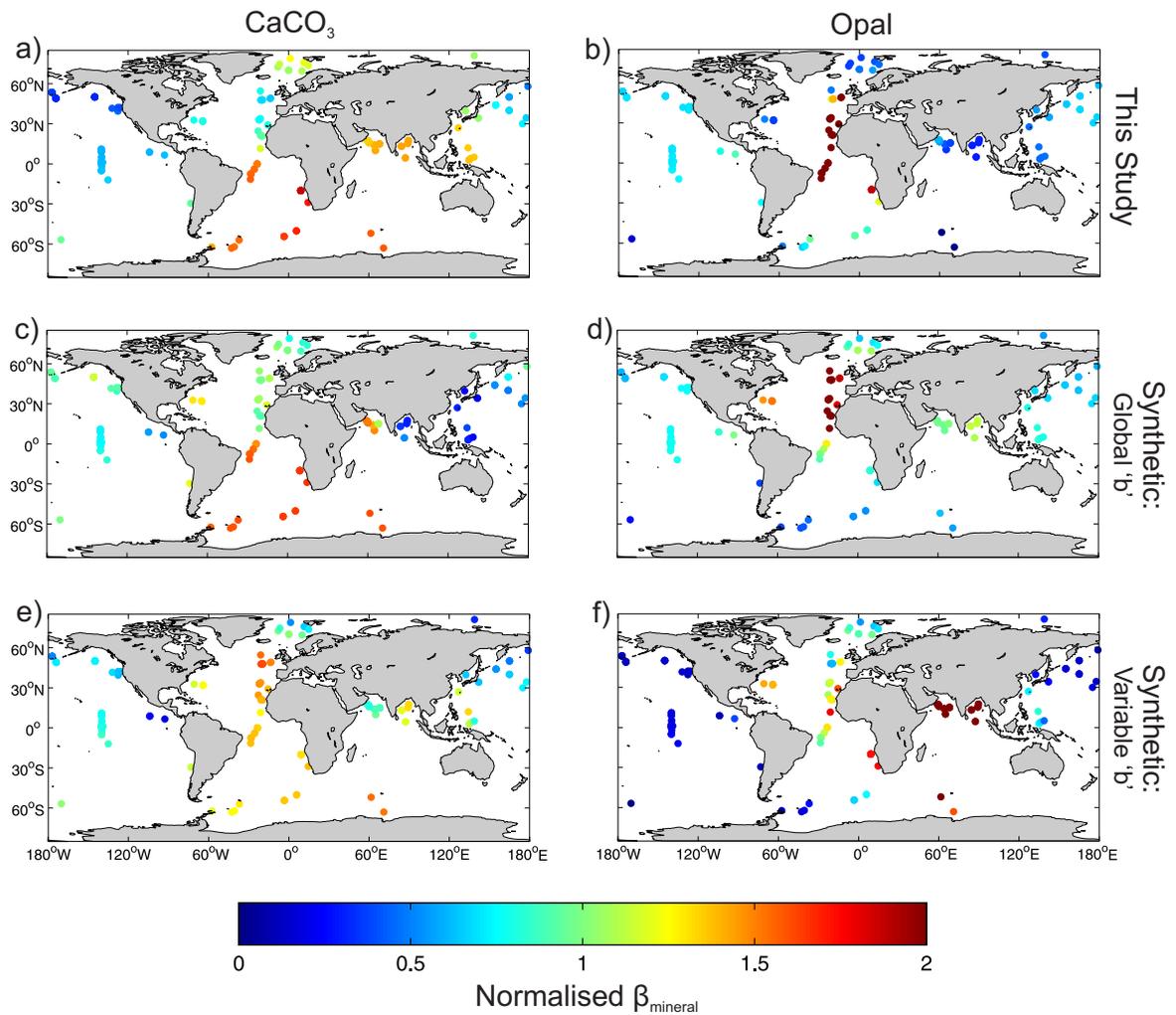
Figure 3.13: The spatial distribution of coefficients calculated using geographically weighted regression on the synthetic datasets shown in Figure 3.12. The GWR analysis uses a bandwidth of 66 nearest neighbours. Panels (a) and (b) are the same panels from Figure 3.7 for comparison. Panels (c) and (d) use the global Martin Curve whilst (e) and (f) use spatially variable Martin Curves. All coefficients are normalised to the mean of each set for ease of comparison.

ballasting mechanism.

3.4   Discussion

Geographically weighted regression (GWR) offers a promising technique to explore the regional variability of regression coefficients. However, its application to sediment trap data is limited both by the amount of data available and the geographical distribution of sampling, which is somewhat clustered, being a collection of individual research projects focused on specific areas and research questions. The overall size of the dataset is also at the lower limit of size to which GWR may be successfully applied (Charlton, M.E., 2012 pers. comm.). GWR is also subject to some of the caveats to previous global mean (and annual average) analyses and which may contribute to the observed variability. For instance, differences in the origin of sinking particles could have a significant influence on the degree of coupling between POC and ballast mineral biasing the use of the ballast hypothesis as a global mechanism. Foraminiferal $CaCO_3$ can be de-coupled from POC due to the reproductive cycle, where the gametes (contributing POC) abandon the organism, leaving particles that are predominantly $CaCO_3$ to sink (*Loubere et al.*, 2007). *Ziveri et al.* (2007) showed that the relationship between POC and $CaCO_3$ in different species of coccolithophores can also be variable, leading to differences in regression coefficients. Similarly, *Thunell et al.* (2007) observed variable POC to opal ratios in different coastal diatom species which contributed to variation in regression coefficients. Because sinking particles captured by sediment traps are expressed as bulk $CaCO_3$ and opal, it is likely this could be a source of variability in the data which cannot be accounted for here. The method of estimating lithogenic fluxes has previously been highlighted as an issue in analysing fluxes (*Boyd and Trull*, 2007). Defined as the difference between total mass flux and the main biogenic components, the lithogenic fraction can be more accurately termed the residual flux (*Salter et al.*, 2010) and may include material that is unaccounted for. Unfortunately this is not something that has

been possible to address in this study. A collation of studies where the sources of any lithogenic material in sediment traps have been explicitly identified, such as *Ternon et al.* (2010), would possibly facilitate future work on this issue.

Despite these caveats, GWR removes some of the subjectivity in choosing subsets of data and provides a framework for dealing with statistical problems, *e.g.,* such as increased variance of coefficients. GWR therefore offers a more objective technique and one able to explore spatial patterns at a finer resolution than the basin-scale analyses previously undertaken.

3.4.1   Why is there spatial variability in carrying coefficients?

If there was a systematic mechanism linking increased sinking velocity of particulate organic matter with $CaCO_3$ or between the efficiency of export to depth and protection from remineralization (*Armstrong et al.*, 2002; *Klaas and Archer*, 2002), one would expect to observe consistent relationships between POC and the presence of minerals at a global scale. However, in the GWR analysis used here significant spatial variability in the carrying coefficients has been identified but with generally coherent large-scale spatial relationships. If models of the ocean carbon cycle are to be improved and hence better quantify feedbacks and potential impacts of global change, it is essential to understand why any coherent patterns occur at all and why regions differ from one another. The underlying complexity and challenge in mechanistic interpretation is illustrated by considering the quantitative importance observed for opal in the Atlantic and $CaCO_3$ in the Southern Ocean, which is somewhat counter-intuitive considering the global distribution of productivity by major plankton groups, with relatively little opal exported from the surface of the tropical Atlantic and relatively little $CaCO_3$ exported in the Southern Ocean. The inference is that fluxes of POC are quantitatively associated with fluxes of opal in the Atlantic despite this being an area of very low opal fluxes. The inverse applies to the Southern Ocean and $CaCO_3$. This suggests that

there is more to the controls of sinking POC than just the dominant mineral phase being produced at the surface.

Multiple linear regression reflects the joint variability between all three minerals and POC and therefore coefficients reflect variability related to the specific combination of all flux components and should not be interpreted in terms of any one mineral component in isolation. *Honjo et al.* (2008) demonstrated that the relationships between the proportions of POC, $CaCO_3$ and opal within mass fluxes are broadly aligned with regional patterns such as biogeochemical provinces. They used this to define broad regions that are dominated either by POC, $CaCO_3$ or opal. The patterns of coefficients in Figure 3.7 define broad regions and appear consistent with this interpretation, being visually comparable to biogeochemical provinces (and which have been demonstrated are not artifacts of the statistical method or inter-basin influence; see Figure 3.10). The regional regression coefficients observed here may then reveal variability derived from the specific combinations of flux components that ultimately derive from specific ecosystem characteristics.

Recent studies have built on this alternative interpretation of the observed relationships between POC and minerals, focusing on the variability originating from ecosystem processes. *Francois et al.* (2002) originally suggested that differences in the biodegradability of POC derived from $CaCO_3$-dominated ecosystems resulted in efficient POC transfer rather than from the direct presence of $CaCO_3$ itself. Developing this methodology further with satellite data, *Henson et al.* (2012) found that the export efficiency of particles from the surface is low in carbonate dominated regions but transfer efficiency of POC to depth is high, with the reverse being true in regions dominated by diatoms. This complements similar findings in the Southern Ocean (*Lam and Bishop*, 2007) and the identification of similar variability through the interaction between the timing and intensity of activity in producer and consumer communities (*Lam et al.*, 2011). Our results are consistent with the relationships between mass flux components being derived from ecosystem characteristics, and support the findings of these studies

in suggesting that the key factor in the variability of POC reaching the deep ocean is the interaction between producer and consumer communities in different ecosystems, possibly reflecting the resulting differences in exported organic matter biodegradability. In ecosystems with highly seasonal productivity, mismatches in the timing of producer and consumer activity can result in labile organic matter being exported which is then readily remineralized in the water column. Likewise, in non-seasonal ecosystems more constant production and consumption can result in the export of relatively refractory material which is more resistant to remineralization in the water column and sinking to depth (*Lam et al.*, 2011). The spatial analysis presented in this chapter suggests that even within diatom and carbonate dominated regions there are distinct variations which may relate to specific differences in this packaging function, such as the distribution of different species and behavior of zooplankton present, and the subsequent action of bacteria (*Lam and Bishop*, 2007; *Buesseler and Boyd*, 2009).

3.4.2   Implications for the ballast hypothesis and modelling

The ballast hypothesis has been highly influential in proposing that fluxes of POC can be mechanistically linked to fluxes of ballast minerals such as $CaCO_3$. The basis of this hypothesis is the observed strong global quantitative relationship observed between POC and $CaCO_3$. However, this strong relationship does not apply to all regions (or even sub-regions) in the ocean and is not constrained to $CaCO_3$ alone. In effect, the strong relationship appears to be an artefact of averaging fluxes on a global scale, masking important regional variability. This could occur because globally there is a large range of flux magnitudes in both POC and $CaCO_3$ which is larger than the variability of the regional POC:$CaCO_3$ ratios. This is evident in Figure 3.5 where different basin scale ratios can be seen within the global data distribution. Using the relative standard deviation (RSD: standard deviation divided by the mean) as a comparable measure of the variation within each dataset, the individual fluxes of $CaCO_3$ and POC have RSD values of 0.62 and 0.71 respectively whereas the RSD

of the CaCO$_3$ carrying coefficients is smaller at 0.45. As such, the appearance of a global correlation between CaCO$_3$ and POC could be exaggerated as a result of considering data at a global scale. Equally, this may bias the interpretation of such relationships without information on spatial variability. In a similar manner, *Lam et al.* (2011) produced global coefficient values when averaging particle concentration data through time but which masked important temporal variability, suggesting a different mechanistic interpretation. The analyses of temporal variability (*Lam et al.*, 2011), transfer efficiencies of POC (*Francois et al.*, 2002; *Henson et al.*, 2012) and the regional variability highlighted in this study all suggest that ecosystem-based mechanisms are influential in setting the efficiency of the biological pump. While this does not rule out ballasting as a mechanism, *e.g., Ploug et al.* (2008); *Engel et al.* (2009a); *Iversen and Ploug* (2010), it shifts the focus away from a simple causal physical explanation.

The ballast hypothesis has inspired specific parameterizations using the global statistical coefficients to be adopted in a number of global ocean carbon cycle models, *e.g.,* HAMOCC5.1 (*Howard et al.*, 2006), CCSR COCO 4.0 (*Oka et al.*, 2008), POTSDAM-C (*Hofmann and Schellnhuber*, 2009), and MEDUSA (*Yool et al.*, 2013) but see also PISCES (*Gehlen et al.*, 2006). The results from this chapter suggest that using these global statistical parameters explicitly in models requires careful reconsideration. However, moving away from a common global mechanism to parameterizations able to capture spatial variability in the flux relationships is not trivial. Application of a simple prescribed 'map' (distribution) of carrying capacities, while potentially improving the simulation of dissolved nutrient and carbon distributions in the ocean interior, is likely to fail to provide an appropriate response to global change. Instead, there needs to be a shift in the focus away from a geochemical-based understanding and parameterization approaches towards a more ecological-based understanding of fluxes, as suggested by (*Ragueneau et al.*, 2006; *Boyd and Trull*, 2007), which will require alternative mechanistic representations in models. Reflecting ecosystem function will then lead to potentially very different feedback mechanisms when considering ocean acidification

and climate change (*Henson et al.*, 2012). However, such a shift towards an ecological-based understanding requires the representation of complex ecosystem processes and an increased need for data for validation, highlighting the utility of empirical parameterisations such as ballasting. A move towards embracing the complexity of processes may add additional uncertainties to our ability to use models to look at past and future changes in the biological pump. In general, further work will be needed to develop flux parameterizations that link aspects of ecosystem function with organic matter particle fluxes. In this respect, models are beginning to consider suspended particles, slow sinking particles that are subject to remineralization in the water column, and fast sinking particles that may be subject to ballasting to facilitate a more complex biological pump in the ocean interior (*Riley et al.*, 2012; *DeVries et al.*, 2014). Considering the range of mechanistic interpretations and inherent uncertainty this introduces, further work is needed to quantitatively constrain the range of feedback processes through more observations of the biological pump and modelling. Given the need for more observations of the biological pump to achieve a better representation with which hypotheses like ballasting can be eveluted against, the next chapter explores a method of estimating particulate organic matter flux curves in the global ocean.

CHAPTER 4

Can organic matter flux profiles be diagnosed using remineralisation rates derived using modelled ocean transport rates?

4.1 Introduction

Sediment trap studies show that the vertical flux of POC can be described empirically by a power-law curve (*Martin et al.*, 1987; see Figure 4.1a) where POC is rapidly remineralised in the upper water column (<1000m) leaving a small fraction (5-10%) of POC that sinks to greater depths. The exponent of the Martin Curve (*b*) reflects the distribution of POC with depth and whether POC is remineralised higher or lower in the water column. Recent studies using global sediment trap observations, ^{234}Th fluxes and particle concentration data suggest a highly heterogeneous pattern of flux profiles and the existence of a general latitudinal trend (*Lutz et al.*, 2007; *Honjo et al.*, 2008; *Henson et al.*, 2012). POC fluxes in high latitude regions decrease faster with depth, *i.e.,* they display shallower flux curves than in low latitude regions where a greater proportion of POC is remineralised at deeper depths (Figure 4.1a). Analysis of the POC fluxes supports a number of mechanisms that explain these spatial patterns including increased sinking velocity of particles via a 'ballast' effect from minerals such as $CaCO_3$

(*Armstrong et al.*, 2002; *Klaas and Archer*, 2002) (see also Chapter 3); aggregation of particles (*Passow*, 2004); and surface ecosystem dynamics, such as the level of nutrient recycling and blooms altering characteristics of the particles being exported (*Lam et al.*, 2011; *Le Moigne et al.*, 2012). However, further progress towards fully understanding which mechanisms are key their representation in models has been hindered by the relatively low sampling density of long-term sediment trap deployments, particularly in regions such as the Southern Ocean, *e.g., Le Moigne et al.* (2012) and Chapter 3.

Understanding the underlying reasons for the spatial patterns in remineralisation is a valuable step in understanding the key mechanisms driving the biological pump. This is key to understanding how the biological pump will respond to both past and current changes in climate (*Honjo et al.*, 2014). Atmospheric CO_2 concentrations have been shown to be sensitive to changes in flux profiles when modelled, primarily via the effect of redistributing DIC in the modern ocean interior (*Kwon et al.*, 2009). Atmospheric CO_2 is also potentially sensitive to changes in the POC flux to deep-sea sediments relative to fluxes of calcium carbonate ($CaCO_3$) over longer timescales (*Archer and Maier-Reimer*, 1994; *Roth et al.*, 2014). These studies have only considered global changes in remineralisation depths and not spatially variable changes that could result in additional uncertainty in the response of CO_2. Another source of uncertainty in these model studies is in the representation of vertical fluxes. Methods range from explicit mechanistic models with sinking and remineralisation rates (*Gehlen et al.*, 2006; *DeVries et al.*, 2014), to the use of empirically derived functions such as the Martin Curve (*Kwon and Primeau*, 2006; *Kwon et al.*, 2009; *Kriest et al.*, 2010).

A potential approach to increasing and enhancing the resolution of POC observations is to use climatological fields of dissolved nutrients to estimate remineralisation rates. Fields of dissolved nutrients, such as from the World Ocean Atlas, offer a much higher resolution of data. The rates of remineralisation are related to vertical flux curves by the fact that the vertical profile of remineralisation rate is the first derivative of the vertical profile of fluxes (Figure 4.1a and 4.1b). Therefore, if organic matter reminer-
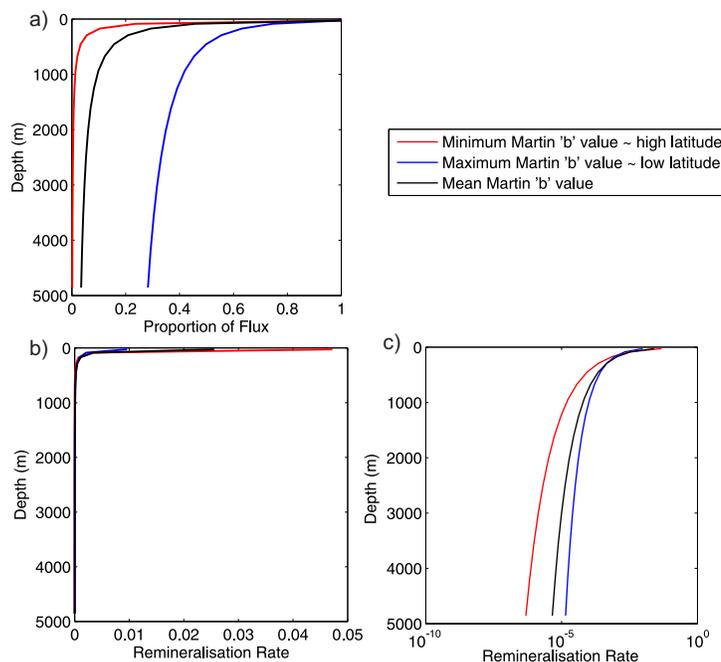
Figure 4.1: The range of observed Martin Curves and associated remineralisation rate profiles (a) The mean (b=-0.639) and global range (b=-1.18 to -0.24) of Martin Curves found by *Henson et al.* (2012) calculated for a unit flux and export depth (z_0=25m) via $F_z = 1 * \frac{z}{z_0}^b$ (b) The first derivative of each flux curve in panel a, equivalent to a vertical profile of organic matter remineralisation calculated as $\frac{dF_z}{dz} = \frac{b}{z_0} * \frac{z}{z_0}^{(b-1)}$.

alisation rates can be estimated, this may provide a way of generating high resolution estimates of flux curves to complement sediment trap analyses. The global distribution of a biological nutrient, such as phosphate (PO_4) or dissolved inorganic carbon (DIC), results from the net action of the biological pump (uptake during photosynthesis at the ocean surface and subsequent remineralisation of organic matter in the interior) in combination with physical processes (*e.g.*, air-sea gas exchange) and other biological processes (*e.g.*, denitrification) that are integrated through time via ocean circulation. These processes need to be taken into account to estimate remineralisation rates.

Apparent Oxygen Utilisation (AOU; in concentration units, *e.g.*, μmol kg[-1]) has a long history of use as a measure of the net organic matter remineralisation in the ocean interior. By assuming that the surface is in equilibrium with the atmosphere with respect to oxygen (O_2), AOU is calculated as the difference between observed [O_2] at any point in the ocean interior and the concentration of O_2 at saturation. Any reduction in

[O_2] indicates the net effect of aerobic respiration during the remineralisation of organic matter. AOU can be converted to apparent oxygen utilisation rates (AOUR; μmol kg^{-1} yr^{-1}) by combing it with tracers that estimate the time the water mass has last been in contact with the surface such as radiocarbon or CFCs. AOURs can be related to carbon remineralisation rates via stoichiometric ratios, *e.g., Jenkins* (1982). *Feely et al.* (2004) compiled AOUR-derived profiles of organic carbon remineralisation rates for 10 regions in the Pacific noting higher rates in the North Pacific relative to the South Pacific and similarities between regions with high rates and higher $CaCO_3$ fluxes. However, because the *in-situ* AOU of a water parcel reflects the history of remineralisation over its whole trajectory in the ocean interior, *e.g.,* along an isopycnal, the AOUR is an average rate that has been shown to be biased towards shallower regions where remineralisation rates are higher (*Stanley et al.*, 2012). Therefore, when relating AOURs back to flux profiles by integrating them vertically, estimates are representative of large oceanic regions only and not amenable to the analysis of spatial variability, *i.e.,* there is a high degree of spatial smoothing (*Sonnerup et al.*, 2013; *Stanley et al.*, 2012).

An alternative to combining AOU and age tracers is to use the spatial gradients in tracers created by circulation. These methods attempt to separate out and quantify the change in tracer concentrations over time at any point in the ocean interior occuring from circulation leaving the change in tracer due to biological processes such as reminerlisation. Gradient based approaches aim to solve for the effect of mixing by defining a water mass as the sum of mass fractions from different sources, *e.g., Anderson and Sarmiento* (1994); *Broecker et al.* (1998); *Gebbie and Huybers* (2010). A recent development of this method, the Total Matrix Intercomparison (TMI) method described by *Gebbie and Huybers* (2010), solves for up to 6 mass fractions for each grid-box in a 4° x 4° resolution with 33 vertical levels using temperature, salinity, $\delta^{18}O$, and nutrients with an additional source term reflecting organic matter remineralisation. The source term is related to the nutrients using stoichiometric ratios. The TMI method therefore produces a high resolution field of remineralisation estimates. However, these terms only

reflect an amount of remineralisation not the rate of remineralisation, *i.e.,* TMI predicts mol PO_4 remineralised in a grid box rather than mol PO_4 yr^{-1} (more specifically TMI predicts the amount of remineralisation over the timescale of circulation, effectively the residence time, in each grid-box. The residence time varies between boxes such that the remineralisation amounts cannot be related directly to eachother). This is because the TMI method models the pathways of ocean transport but not the rates of transport (*Gebbie and Huybers*, 2010).

Spatial gradients in tracers have also been used to diagnose export fluxes of calcium carbonate (*Sarmiento et al.*, 2002) and opal (*Sarmiento et al.*, 2004a). The method is based on taking the ratio between the concentrations of alkalinity and nitrate between the upper 100m of the ocean and the 100m below. The ratio is assumed to reflect the strength of $CaCO_3$ and organic carbon export. The method assumes a dominance of vertical gradients (*Sarmiento et al.*, 2002). However, a model-based assessment of the method suggests that biases could occur due to the remineralisation of dissolved organic matter (DOM) and strong meridional transports that violate the assumption of processes occurring primarily in the vertical dimension (*Jin et al.*, 2006). Although this method is not applicable to estimating remineralisation rates in the ocean interior (as tracer gradients in the ocean interior cannot be assumed to predominantly vertical), these studies highlight that the assumption that all remineralisation originates from particles that have sunk through the water column also requires consideration.

4.1.1   Aims and Outline

Ocean circulation models offer the opportunity to estimate remineralisation rates of organic matter from tracer data by exploiting the calculated modelled transport rates to estimate the contribution of circulation in a similar process as the TMI method described above. The aim of this chapter is to explore the feasibility of inferring flux profiles of particulate organic matter from remineralisation rates that have been de-

rived from observed tracers using modelled ocean transport rates. First a method of estimating remineralisation rates using modelled circulation in form of transport matrices is first introduced in section 4.2 with an illustrative example. A set of model experiments are used as a synthetic dataset with which to test the sensitivity of the approach to various sources of error. Section 4.5 explores the potential uncertainties associated with the assumptions when inferring flux curves from remineralisation rates. Finally, the potential future directions for the method are discussed.

4.2 Estimating remineralisation rates using modelled ocean transport rates

4.2.1 Concept and Background

Remineralisation rates (q) can be calculated as the amount of tracer needed to maintain tracer observations at steady state once the effects of ocean circulation have been accounted for, $i.e.,$ transport divergence ($Deutsch\ et\ al.$, 2007). This requires quantifying the total flux of a tracer into a grid box from circulation ($\sum_{i=N}^{N} F_i^{in}$) from N neighbouring locations (i) and the flux out ($F^{out}C$) at a given location in the ocean interior (Equation 4.1):

$$\frac{dC}{dt} = \sum_{i=1}^{N} F_i^{in}C_i - F^{out}C + q \tag{4.1}$$

The ocean circulation fluxes (F_i^{in} and F_i^{out}) can be estimated from ocean circulation models. A practical method for this approach is to apply a transport matrix approach. A transport matrix, hereafter abbreviated to TM, is a representation of transport rates in the form of a sparse matrix that is derived empirically from an ocean circulation model ($Khatiwala\ et\ al.$, 2005; $Khatiwala$, 2007). Using a TM is akin to the fixed transport rates implied in a box model ($e.g.,$ 'LOSCAR': $Zeebe$, 2012) but with the

advantage that the rates are diagnosed from a dynamic ocean model providing a realistic representation of ocean circulation with a much higher spatial resolution.

For every grid-box in the model the TM defines the relative fluxes in and out of a tracer for every grid-box due to ocean circulation during a single time-step of the model (see Table 4.1 for an illustrative example). Using the TM in place of the model circulation, the evolution of a tracer in time is then:

$$\mathbf{c}^{n+1} = \mathbf{A}\mathbf{c}^n + \mathbf{q} \tag{4.2}$$

where $\mathbf{A}$ is the TM (unitless), $\mathbf{c}$ is a vector representation of the gridded three-dimensional tracer field (mol kg^{-1}) with the superscript referring to the time-step index. $\mathbf{q}$ is a vector representation of any non-circulation related sources and/or sinks for a tracer over the timestep of the model (mol kg^{-1}), for instance, due to remineralisation. Note that the bold $\mathbf{q}$ indicates a vector of all grid-boxes in comparison to a single grid-box in equation 4.1. Assuming steady state (where $\mathbf{c} = \mathbf{c}^{n+1} = \mathbf{c}^n$), the simplest approach to estimating remineralisation rates using equation 4.2 is to solve for $\mathbf{q}$ directly, given the TM and a steady state tracer:

$$\mathbf{q} = -(\mathbf{A} - \mathbf{I})\mathbf{c} \tag{4.3}$$

where $\mathbf{I}$ is the identity matrix (see Table 4.1 for the role of $\mathbf{I}$ and Appendix B for derivations of the TM equations). Applying equation 4.3 with a steady-state tracer field is the same as initialising an ocean circulation model with observed tracer concentrations and then stepping the model forward for one time step. $\mathbf{q}$ will therefore reflect the exact interior source/sink terms needed in one time step to maintain the steady state tracer concentrations given the model transport rates. Estimates of remineralisation rates are found in $\mathbf{q}$ but are referred to here as interior source/sinks (ISS) as they do necessarily represent a mechanistic process.

Table 4.1: Example of using a transport matrix to calculate PO_4 remineralisation (μmol kg^{-1} dt^{-1}) in one grid-box from PO_4 concentrations (μmol kg^{-1}) given in **c**. Grid-boxes are arbitrarily numbered, where the 1 is the grid-box where the calculation is taking place. Descriptions in brackets give an example of the relationship of the neighbouring grid-boxes to the central grid-box. Coefficients from **A** represent the change in a tracer due to circulation after a single time-step, *e.g.*, the proportion of tracer concentration left in grid-box 1 after one time step is 0.9816 (see equation 4.2). Coefficients from (**A-I**) are the same except now grid-box 1 is equivalent to 0.9816-1.0000 (see equation 4.3). The sum of the coefficients are shown underneath. The amount in bold at the bottom of the table is the estimated remineralisation combining (**A-I**) and **c** (**q** in equation 4.3).

Grid-Box	A	(A-I)	c
1 ('flux out')	0.9816	-0.0184	2.3439
2 ('north')	-0.0007	-0.0007	2.3430
3 ('south')	0.0086	0.0086	2.4334
4 ('east')	0.0002	0.0002	2.3529
5 ('west')	0.0005	0.0005	2.3615
6 ('up')	0.0097	0.0097	2.4433
7 ('down')	0.0001	0.0001	2.3369
	1.0000	0.0000	**0.0011**

4.2.2   Example using a General Circulation Model Transport Matrix

As an example of the approach, the annual average TM derived from a 2.8° configuration of the MITGCM model (available online: `http://www.ldeo.columbia.edu/~spk/`) is used to invert a regridded annual climatogical PO_4 field (World Ocean Atlas 2009; *Garcia et al.*, 2010), using equation 4.3 (Figure 4.2). The resulting ISS estimates at shallower depths of the ocean interior (85m) show some spatial patterns that are consistent with expectations of general export patterns, such as higher rates in the equatorial upwelling regions, and the subpolar regions of the Southern Ocean (*Henson et al.*, 2011) (Figure 4.2a). However, negative ISSs also exist, indicating a sink of PO_4, such as in regions of the subpolar Southern Ocean and Pacific equatorial upwelling. At deeper depths (2030m) the elevated values in the Southern Ocean match the shallower pattern, but overall there are fewer clear spatial features and a more random pattern of positive and negative ISSs (Figure 4.2b). The existence of so many negative values gives rise to near-zero values when averaging over large spatial scales.
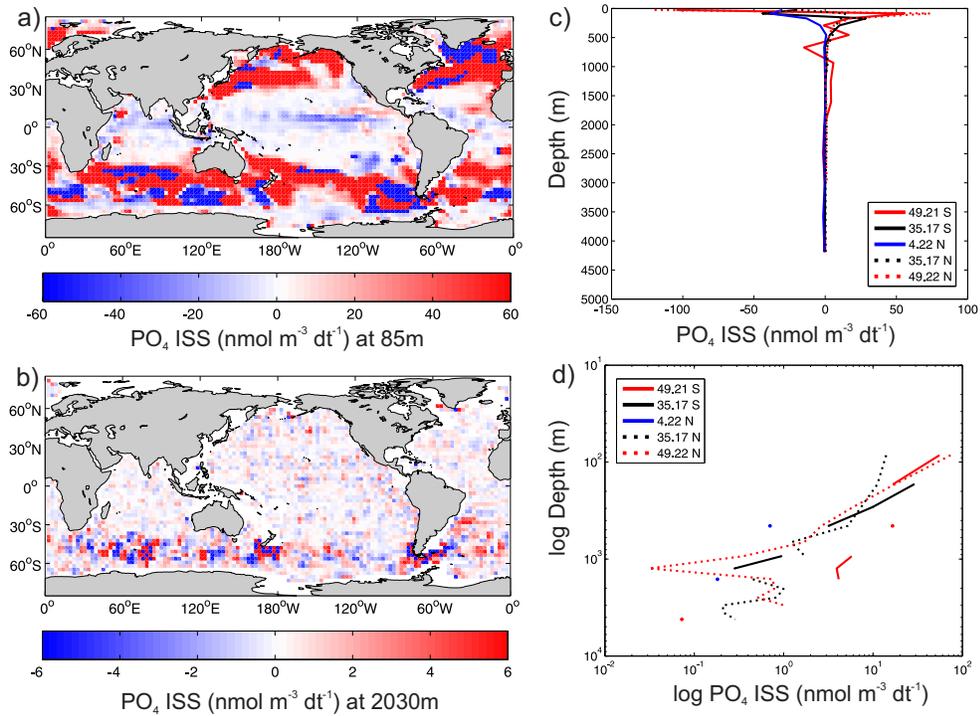
Figure 4.2: Example of using a GCM transport matrix to estimate PO_4 remineralisation rates. (a) The estimated PO_4 remineralisation rates generated using the MIT GCM transport matrix at 85m and (b) 2030m. (c) Vertical profiles of PO_4 remineralisation rates estimated using transport rates from an MITGCM transport matrix (equivalent to Figure 4.1b) are shown from latitudes corresponding to equivalent regions in the Pacific at 223°E. and (d) on a log scale, comparable to Figure 4.1c.

Vertical profiles of the PO_4 ISSs show a range of features (Figure 4.2c). Several show negative values at the surface, which is expected given this will reflect uptake of $[PO_4]$ by phytoplankton during photosynthesis. The Pacific profile at 35.17°S is smooth and fitting expectations of a remineralisation curve in Figures 4.1c and Figure 4.1d. The profile at 4.22°N however shows increasingly negative ISSs at shallower depths. This example shows that a simple inversion of PO_4 observations using this approach is susceptable to large errors that will likely hinder their interpretation. The difference between the model circulation and observed circulation is a clear reason why there are large errors. In the following section, a method is developed to quantify and characterise these errors.

4.3   Methods

4.3.1   Model Description

To explore the errors when using modelled transport rates, a synthetic dataset is derived using the Earth System model 'GENIE' (*Ridgwell et al.*, 2007a) (see also Chapter 2). GENIE features a 3-D ocean circulation model coupled to a 2-D energy-moisture balance model of the atmosphere and a dynamic-thermodynamic sea-ice model (see *Edwards and Marsh*, 2005). In the configuration used here, the ocean model is non-seasonally forced and solved on a 36x36 equal area horizontal grid (10° longitude by 3°- 15° latitude) and 8 vertical layers. The biogeochemical model is that described in *Ridgwell et al.* (2007a) with one exception described below.

Our choice of GENIE over other possible models and available transport matrices reflects a number of considerations. The configuration of GENIE used here was derived using an ensemble, where combinations of parameters relevant to the physical circulation, such as isopycnal diffusivity, were sampled to test the sensitivity of the model circulation to parameter values and find an optimal set of parameters (*Annan et al.*, 2005). The availability of this ensemble provides 54 plausible estimates of the modern ocean circulation enabling an assessment of the errors associated with uncertainty in model circulation states (see Figure C.1). The annual average circulation, coarse resolution and integration speed of GENIE also facilitates the relatively easy and fast retrieval of multiple TMs. This is the first time that transport matrices have been constructed from an Earth system model. The relative simplicity of GENIE also keeps a focus on the methodological concept. For example, the relative simplicity of GENIE means that a simulation can be run for long enough to achieve a steady-state circulation which is not always possible with higher resolution models, which would add an additional uncertainty to any results. It is worth noting that much of this could also be achieved by 'coarse-graining' a TM derived from a higher resolution model, described

in *Khatiwala* (2007), but without the availability of multiple estimates of the modern ocean circulation.

4.3.2   Diagnosing Transport Matrices in GENIE

The method of *Khatiwala et al.* (2005) is adapted to diagnose the model ocean circulation in GENIE. The nth ocean grid-box in GENIE (of a total of 6210 ocean grid-boxes in this configuration) is dyed with $1 \, \mathrm{mol \, kg^{-1}}$ of an inert 'colour' tracer, *i.e.,* a conservative tracer that is only changed by the circulation model. The model is integrated for one time-step and the resulting pattern of the tracer is recovered, vectorised, and forms the nth column of a 6210x6210 sparse matrix. This is repeated for all 6210 ocean grid-boxes in GENIE. Each column of the matrix contains the concentrations of tracer after the model has calculated net ocean transport one timestep for each grid-box in turn. The method of diagnosing the TM in GENIE differs from that detailed by *Khatiwala et al.* (2005) in two ways. First, the approach in GENIE does not use smoother basis vectors and instead uses the simpler method of initialising a single grid-box. Secondly, each grid-box is initialised only once and there is no averaging because the circulation in the 8-level version of GENIE is non-seasonally forced. The circulation can therefore be diagnosed during a single continuous simulation. The corresponding Fortran code for diagnosing a TM in GENIE, its application, and relation to the method of *Khatiwala et al.* (2005) can be found in Appendix B.

The transport matrix algorithm is applied to a simulation in GENIE after 10000 years. This is to ensure that the circulation diagnosed is from a steady-state circulation. After 10000 years deep mean salinity and temperature, as well as the overturning circulation, have all reached steady state suggesting the circulation is at steady state (Figure C.2). The deviations in mean temperature and salinity over the last 1000 years of the simulation are $<1\mathrm{x}10^{-3}$ °C, $<1\mathrm{x}10^{-5}$ PSU respectively. All TMs diagnosed in this chapter are produced in the same way.

In the ocean model component, C-GOLDSTEIN, grid-box volumes are fixed and net precipitation-minus-evaporation at the ocean surface is implemented as a virtual salinity flux rather than a loss or gain of freshwater. To account for dilution/concentration effects in GENIE, the biogeochemical tracer concentrations are salinity normalised before the calculation of ocean transport and converted back to concentrations after. A rescaling of the concentrations is performed to ensure the mass conservation of tracers. The TM therefore requires the salinity normalisation of biogeochemical tracer concentrations:

$$\mathbf{q} = -\mathbf{s} \circ ((\mathbf{A} - \mathbf{I})(\mathbf{c} \circ \mathbf{s}^{-1})) \qquad (4.4)$$

where $\mathbf{s}$ is the vector of salinity from the corresponding model run and $\circ$ indicates the element-wise multiplication of vectors.

4.3.3   Experiment Design

The biogeochemical model described in *Ridgwell et al.* (2007a) with the biogeochemical parameter values described in *Ridgwell et al.* (2007b) is used to produce a synthetic dataset of tracers. The estimated remineralisation rates from a synthetic dataset can be directly compared to the model input providing a exact test of the TM inversion methodology. In this, nutrients are utilised by biological activity in the surface ocean grid-boxes based on a nutrient and light limited scheme. A fraction of the uptake is exported from the surface as DOM which is advected and remineralised. The remaining fraction is exported (34%) as particulate organic matter (POM) which remineralises instantaneously at depth according to the Martin Curve with a global b value of -0.858. The use of a Martin Curve keeps the discussion relevant to its use with sediment trap data and is the only difference between our model set-up and that of *Ridgwell et al.* (2007a,b). POM remaining in the deepest grid-box is completely remineralised to maintain a closed system, *i.e.*, there are no sediment interactions.

Table 4.2: Simulations and Experiments used in this Chapter.

Name	Description
Synthetic Datasets	
SYN	A synthetic [PO$_4$] dataset derived from a 10000 year spin-up of GENIE. A corresponding transport matrix is diagnosed.
SYN-NODOM	As the SYN dataset but with the fraction of DOM exported set to zero.
Experiments	
TWIN	The synthetic [PO$_4$] field is inverted using the TM and compared to the modelled [PO$_4$] remineralisation from SYN
ERR-OBS	The synthetic [PO$_4$] is perturbed with error estimates from World Ocean Atlas observations
ERR-CIRC	The synthetic [PO$_4$] is inverted using 54 TMs diagnosed from an ensemble of modern ocean estimates.
ERR-DOM	Particulate flux curves are estimated from the PO$_4$ remineralisation rates from the SYN and SYN-NODOM datasets.

4.3.3.1   Synthetic Datasets

[PO$_4$] is used as the tracer for inversion by the TM. The synthetic tracer is used as a surrogate for observations for which the associated remineralisation and circulation rates are known explicitly. An alternative tracer for this could be AOU (related to [PO$_4$] via stoichiometric ratios), as this tracks only organic matter remineralisation (regenerated [PO$_4$]) whilst total [PO$_4$] also includes [PO$_4$] that has been advected from the surface (preformed [PO$_4$]). However, AOU derived from observed oxygen concentrations is subject to assumptions about oxygen saturation at the surface (*Ito et al.*, 2004; *Dietze and Oschlies*, 2005; *Duteil et al.*, 2013). To focus on errors deriving only from the method of inverting tracers, PO$_4$ is chosen over AOU. The TM derived estimates of PO$_4$ remineralisation are compared with the total remineralisation of PO$_4$ in each grid-box as diagnosed within the experiment run as mol PO$_4$ kg^{-1} year^{-1} and is converted to

dt^{-1} by dividing by the number of timesteps per year (96; dt=0.01 year). The [PO$_4$] field
is the annual average taken from the last year of a 10000 year spin-up. A corresponding
TM is diagnosed at the end of the 10000 years run. This synthetic dataset is referred
to as SYN (Table 4.2). A second synthetic dataset (SYN-NODOM) is also produced
where no DOM is produced to explore the effect of DOM remineralisation (or more
generally an additional source of remineralisation that is not directly related to sinking
particles) when inferring flux curves from remineralisation rates.

4.3.3.2   Experiments

A number of experiments are designed to explore the sensitivity of the approach to
various sources of error (experiment names are indicated in brackets and also described
in Table 4.2):

1. (TWIN)The TM corresponding to the synthetic dataset (SYN) is used to estimate
   remineralisation rates from the corresponding [PO$_4$] field as a proof of concept of
   the method (a twin-test).

2. (ERR-OBS) Errors associated with the tracer observations are simulated by cal-
   culating 100 random perturbations to the synthetic PO$_4$ concentrations (SYN)
   within one standard deviation. The standard deviation for each grid-box is taken
   as the standard deviation of the WOA [PO$_4$] observations regridded to the GENIE
   grid (*Garcia et al.*, 2010). The standard deviation therefore reflects the variability
   of observed [PO$_4$] in that grid-box.

3. (ERR-CIRC) To explore the effect of circulation uncertainty, 54 individual TMs
   are diagnosed from an existing ensemble of modern ocean circulation estimates
   (*Annan et al.*, 2005). The ensemble is the result of tuning circulation parameters
   to fit modern temperature and salinity fields using an ensemble Kalman filter.
   Each ensemble member is spun-up for 10000 years after which the TM is diagnosed
   and used to invert the synthetic [PO$_4$] field (SYN). The circulation parameters

in *Ridgwell et al.* (2007a); *Ridgwell and Hargreaves* (2007) are an average of the ensemble parameters, such that the 54 TMs offer a range of different but plausable modern ocean circulation estimates. Details of the ensemble parameters and comparisons against the standard configuration can be found in Appendix C: Table C.2.

4. (ERR-DOM) The effect of DOM remineralisation when inferring particulate flux curves from remineralisation rates is explored. As DOM remineralisation can occur in a different water column than where DOM was produced, this is a test of the assumption that reminerlisation rates can be assumed to reflect vertical processes only. As a comparison to the synthetic dataset, an identical experiment but with no DOM created (SYN-NODOM) is created, *i.e.,* all PO_4 is exported as particulate organic matter.

4.4 The Transport Matrix Inversion Method

4.4.1 Assessment of the Transport Matrix Inversion Method

The output from a standard configuration of GENIE is used as a synthetic dataset from which to assess the transport matrix inversion method and identify the sources and nature of the errors involved. Figure 4.3a and 4.3b show the $[PO_4]$ field in GENIE at two depths, directly below the surface (290m) and in the deep ocean (2106m), with the corresponding annual average input of PO_4 from the continual remineralisation of sinking particles and dissolved organic matter at the same depths (Figures 4.3c and 4.3d). The higher remineralisation values calculated for single grid-boxes occur where the remaining particulate flux is remineralised at the seafloor to maintain a closed system, *i.e.,* to ensure there are no losses to sediments (Figure 4.3d). The inventory of phosphorus in DOM integrated over the ocean interior below 175m is also shown (Figure 4.3e). DOM has a visually similar pattern to shallow PO_4 remineralisation, because

they are both linked by export production. The close resemblance and magnitude of the DOM at shallow depths may have implications for using remineralisation rates to infer remineralisation fluxes which will be discussed in section 4.5. Finally, there a correspondence between areas in the Southern Ocean with a large magnitude of remineralisation at shallow depths and locations where convection occurs in the model (Figure 4.3f). This is consistent with high productivity driven by nutrients delivered to the surface via strong vertical mixing.

4.4.1.1 *Twin-test*

To demonstrate and test the method described in section 4.2, the model generated synthetic [PO$_4$] field is first inverted using the corresponding TM as per equation 4.4 (Figure 4.4a). The interior source/sink term (ISS), calculated by inverting the synthetic [PO$_4$] field, is consistent with the model calculated remineralisation with minor deviations from the 1:1 ratio line (Figure 4.4a). This demonstrates the success of the approach as the errors in the synthetic tracer field and circulation scheme are effectively reduced to near-zero in this example (to an order of magnitude of 1×10^{-3} nmol kg^{-1}). The errors between estimated and modelled remineralisation rates cluster around zero with a median proportion of error of 6.3x10^{-5} (Figure 4.4b). The errors from the inversion are randomly distributed at both 290m (Figure 4.4e) and at depth (Figure 4.4f), supporting the suggestion that the SYN run is in steady state (Figure C.2) and that the method to diagnose the TM is producing an accurate representation of the model circulation. Overall, the TM inversion of the synthetic [PO$_4$] estimates the PO$_4$ remineralisation rates very well demonstrating that, in theory at least, modelled circulation rates in the form of a transport matrix can be used to successfully estimate remineralisation rates from a steady state tracer field.
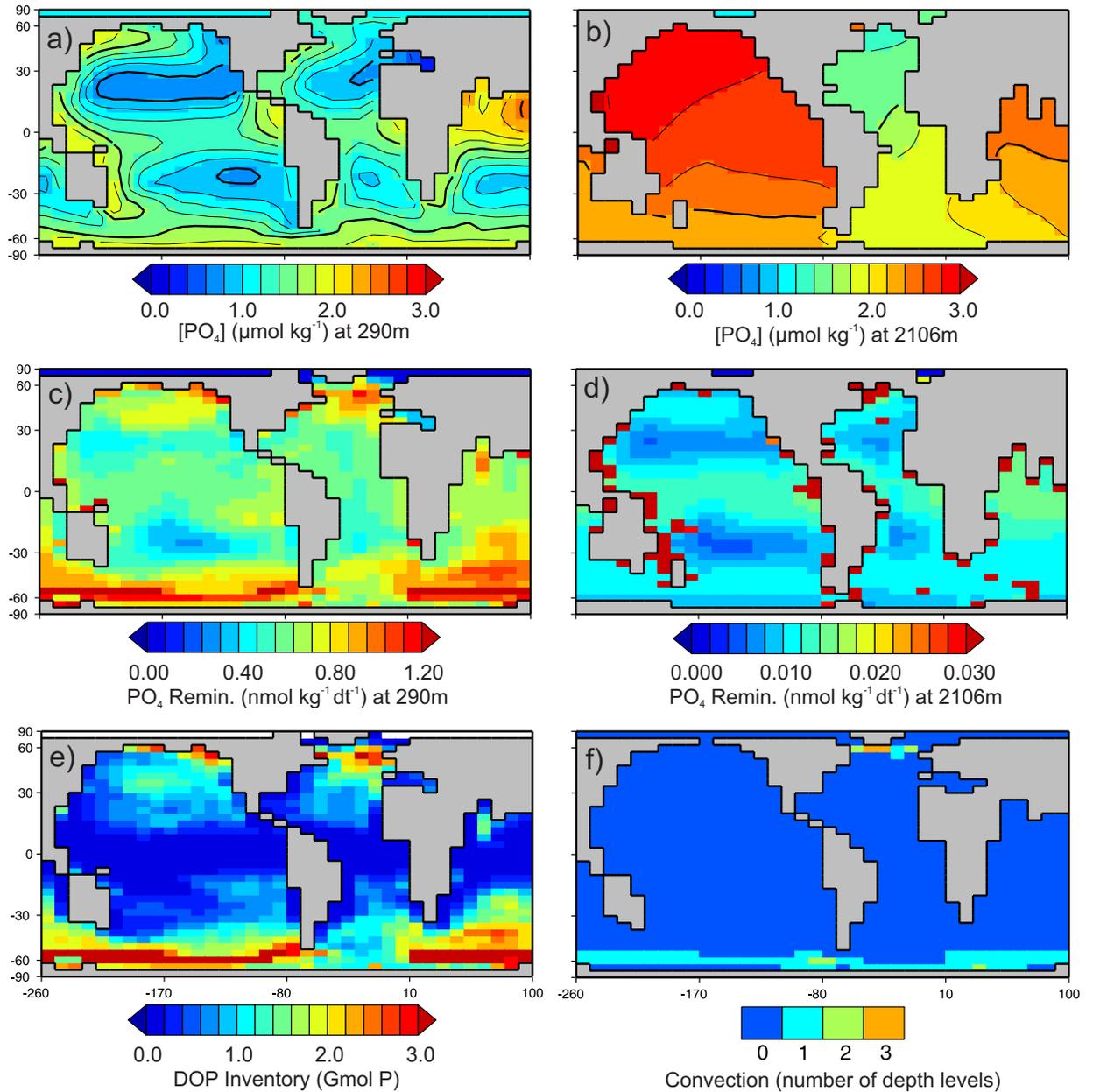
Figure 4.3: The synthetic tracer dataset used for transport matrix inversions. (a) [PO$_4$] (μmol kg^{-1}) at 290m and (b) 2106m. (c) the total annual remineralisation flux of PO$_4$ (nmol kg^{-1} dt^{-1}) at 290m and (d) 2106m. (e) the water column integrated inventory of phosphorus in dissolved organic matter in the ocean interior (mol P). (f) the average number of vertical levels that are involved with convection occurring in a water column (number of depth levels). Higher values indicate deeper and stronger convection in the model.
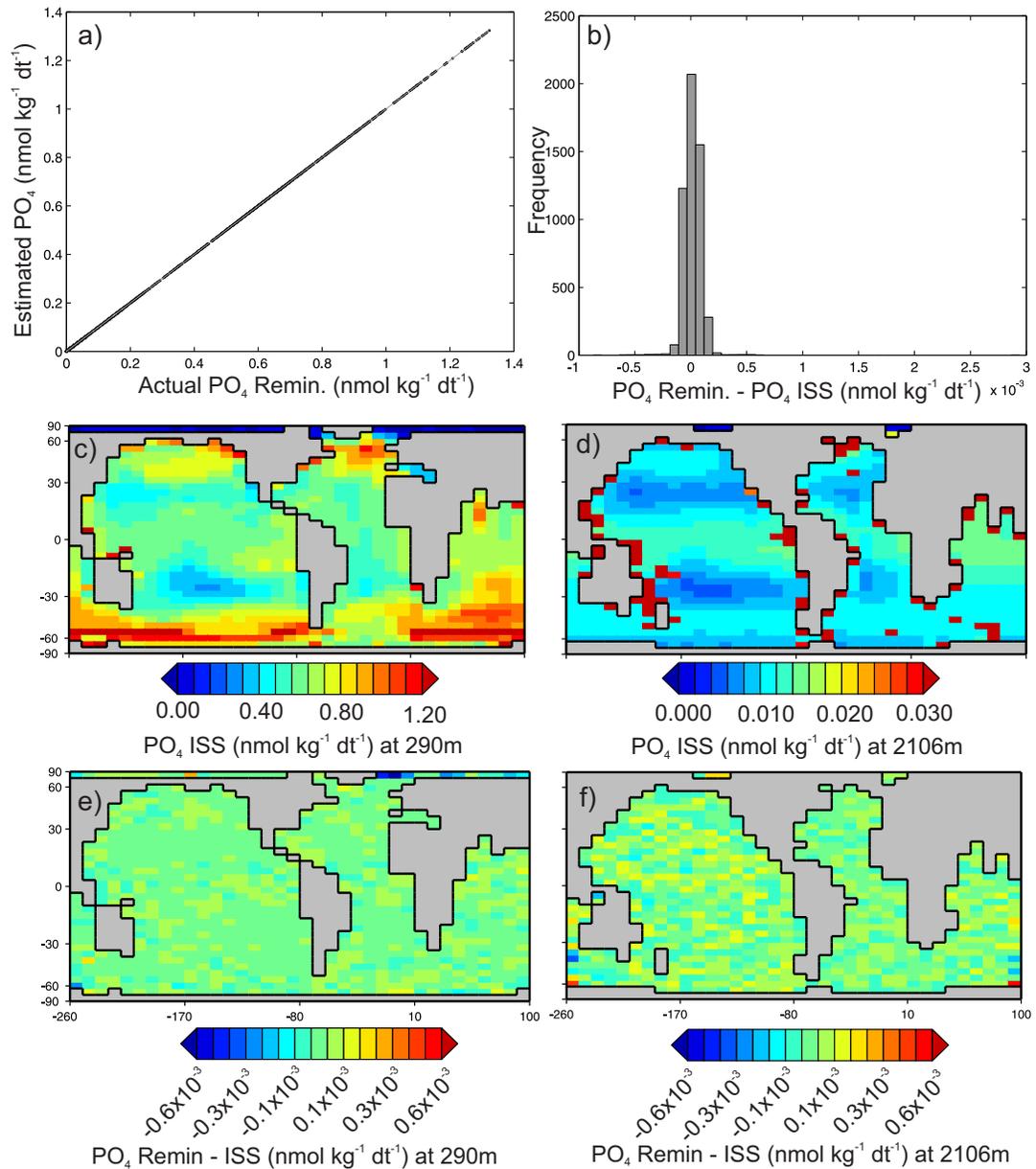
Figure 4.4: Results from inverting the synthetic dataset with its corresponding transport matrix. (a) The interior source/sink estimate for PO_4 estimated when inverting the synthetic $[PO_4]$ field with the corresponding transport matrix plotted against the synthetic PO_4 remineralisation data, *i.e.,* the original input. A 1:1 ratio line is also shown, (b) the distribution of errors for the PO_4 interior source/sink estimates (50 bins sized 0.15×10^{-5}). (c) the interior source/sink estimate for PO_4 at 290m and (d) 2106m. (e) Difference between the inverse interior source/sink estimates and the synthetic remineralisation field at 290m and (f) 2106m.

4.4.2   Sensitivity of Inversions to Sources of Errors

Although remineralisation rates can be estimated by applying transport rates to a tracer field as shown above, there are several assumptions that will introduce error when this is applied to observations. In the following sections, the results of experiments designed to explore these sources of error are described.

4.4.2.1   Error from Observations

Error related to the 1° x 1° World Ocean Atlas annual mean climatology (*Garcia et al.*, 2010) will introduce some uncertainty in the TM inversion due to measurement errors and biases in the climatology itself as well as re-gridding the observations onto a model grid such as GENIE or MITGCM. As a measure of how sensitive the TM inversion method is to these errors, the standard deviation of annual $[PO_4]$ observations in the World Ocean Atlas are regridded onto the GENIE grid (Figure 4.5). The standard deviations in each grid-box are used to produce an illustrative estimate of the uncertainty inherent in the observations. The standard deviations are highest in the coastal regions and at shallower depths (Figure 4.5a *cf.* Figure 4.5b) reflecting where $[PO_4]$ measurements are most variable.

The standard deviations are then used to produce 100 versions of the synthetic $[PO_4]$ field that have been randomly perturbed within the observation uncertainty which are then inverted using the TM. The resulting standard deviation of the PO_4 ISSs are relatively large compared to the ISS values themselves, around 1-3 orders of magnitude larger than the ISS values (Figure 4.5c *cf.* Figure 4.3a). There are positive linear trends between the standard deviation of observations in each grid-box and the standard deviation of the 100 corresponding ISS estimates in each grid box (Figure 4.5c). Grid-boxes with higher uncertainty in the observation results in greater uncertainty in the ISS estimates. However, two distributions can be broadly defined in Figure 4.5c both with separate linear trends that correspond well with the size of the 'flux out' term
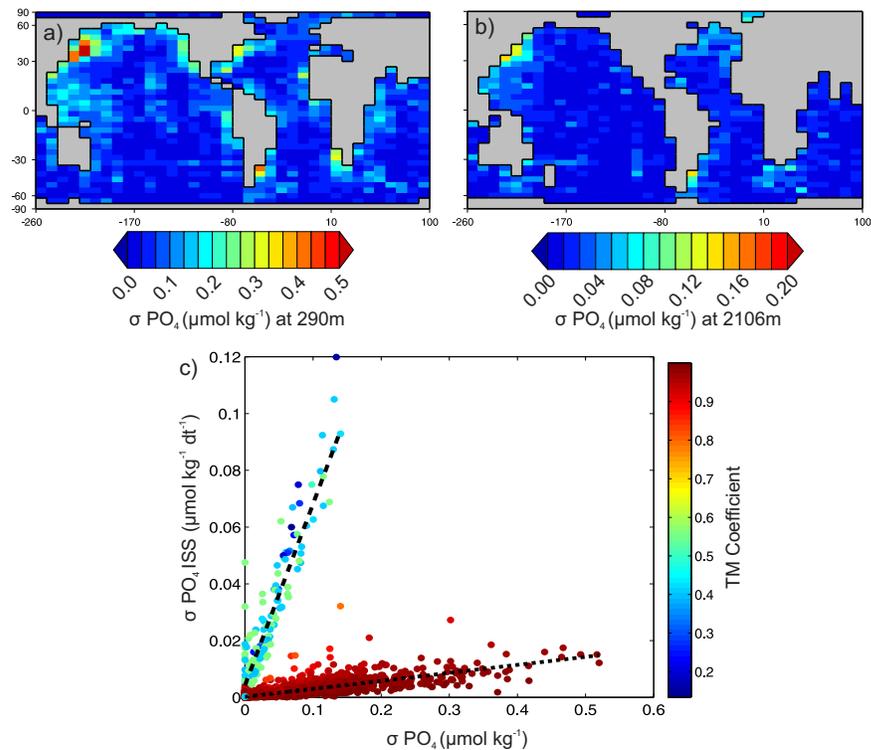
Figure 4.5: Assessment of the errors arising from the uncertainty in [PO$_4$] observations. (a) the standard deviation of [PO$_4$] from the World Ocean Atlas 2009 (*Garcia et al.*, 2010) 1°climatology regridded to the GENIE grid at 290m and (b) 2106m. (c) Standard deviation of all PO$_4$ interior source/sink estimates when the synthetic [PO$_4$] field is randomly perturbed within a normal distribution given by the standard deviation of observations. The two distributions are distinguished by the value of the TM coefficient in the same grid-box shown by the colour bar. A linear regression trend line fitted to data where the coefficient is >0.8 (dotted line) gives a slope of 0.03 (R^2=0.71). When fitted to data <0.8 (dashed line), the slope is 0.64 (R^2=0.84). Note the change in unit size from Figure 4.3c and d.

of the TM (see Table 4.1). Where the central grid-box coefficient is smaller in the TM, leading to correspondingly larger value when used in equation 4.3 (*e.g.*, $\mathbf{A} - \mathbf{I}$), the uncertainty in ISS arising from the uncertainty in the observations is much more sensitive. This suggests that the ISS uncertainty is also a function of the TM itself, *i.e.*, a function of the coefficients within the TM due to the linear nature of the calculation in equation 4.4. The source of uncertainty from observations is therefore higher in the high latitudes consistent with where the flux out term in the TM (see Table 4.1) is largest where convection occurs because this is where the largest transport fluxes are in the model. This is a caveat of the TM calculation rather than from reflecting the

uncertainty of the observations themselves.

4.4.2.2   Error from Circulation Estimates

Another potential source of error when inverting nutrient observations arises from the use of a modelled circulation field that will inevitably have a somewhat poor representation of the circulation of the real ocean. Using a ensemble of modern circulation estimates to invert the synthetic dataset, allows one to explore the effect of errors arising from uncertainties in circulation rates only. As the ensemble reflects a range of varied estimates for the modern ocean circulation, *e.g.,* Table C.1, this is used as a surrogate for model circulations that may be close to but ultimately not matching observed circulation rates. Figure 4.6a shows the mean and 1 standard deviation of the ISS estimates generated when inverting the synthetic [PO$_4$] field with all 54 ensemble TMs. The values furthest from the 1:1 ratio line and those with the largest error bars are located in regions where convection occurs in the model (Figure 4.6a and 4.6c). The strength of overturning varies within the ensemble (Table C.1) suggesting that this structural uncertainty in the model is a likely cause for the wide range of remineralisation estimates. An additional issue is that unlike most grid-boxes in the TM, where the spread of a tracer over one timestep is limited to neighbouring grid-boxes, convection increases the number of grid-box connections in the vertical. This could have the effect of increasing the range of remineralisation estimates because there are more grid-boxes. Figure 4.6b shows the same as Figure 4.6a but with the convection-related values removed. Even in areas where there is no convection in the model, the range of ISS estimates from circulation uncertainty is larger relative to the range of remineralisation values. The range of errors arising from using different circulations are also larger at shallower depths compared to deeper depths in the water column (Figures 4.6c and 4.6d).

To understand why different circulation estimates can have a large impact of ISSs, the size of the PO$_4$ remineralisation flux in a grid box relative to the size of PO$_4$ flux from
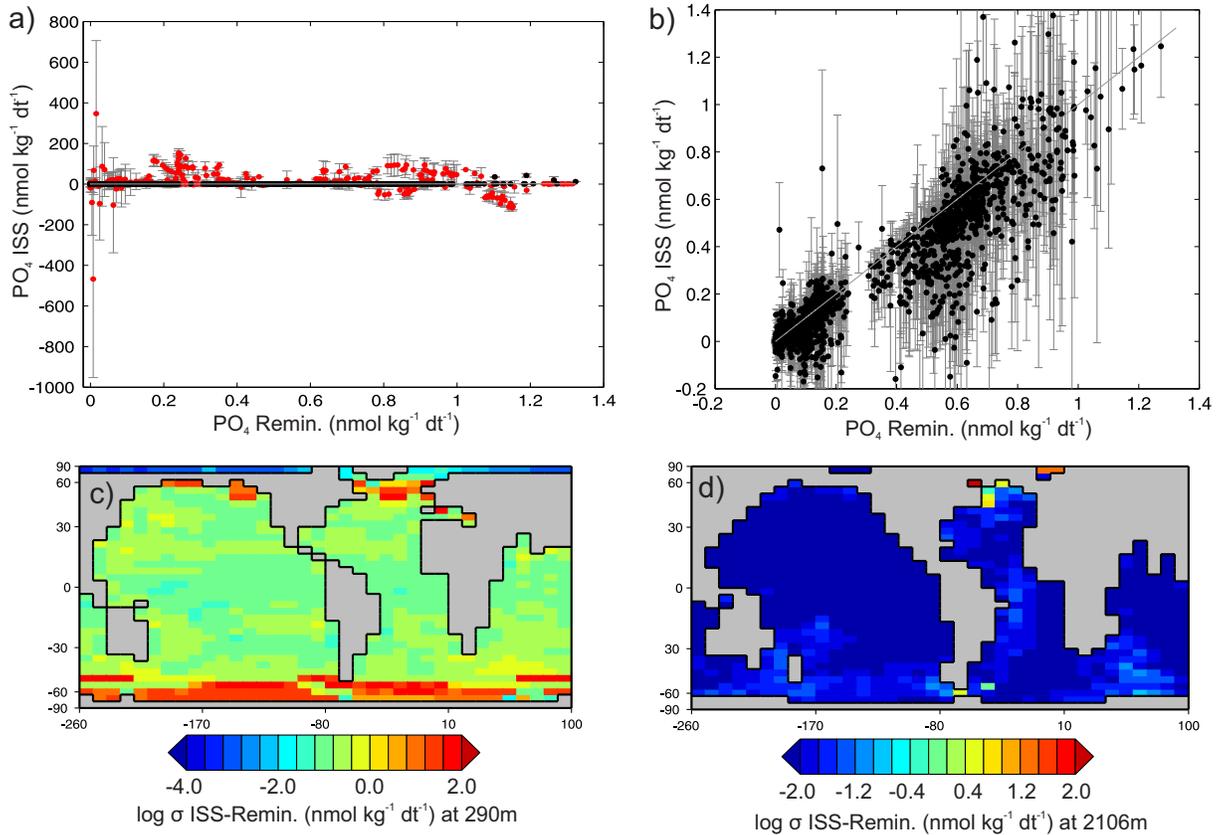
Figure 4.6: Assessment of error arising from using circulation estimates. a) Comparison of ISS estimates for each grid-box from the 54 ensemble members against the synthetic dataset remineralisation. Error bars are 1 standard deviation around the mean. Red values indicate regions with convection. b) As panel a but with the red values removed. Note the difference in scale. c) The standard deviation of PO_4 ISS errors (Model ISS) in μmol kg^{-1} dt-1 for 290m. Values are shown on a log scale. d) As panel c but at 2106m.

the modelled circulation is explored. To illustrate this, the steady state circulation flux of PO_4 into a grid-box is compared with the remineralisation flux of PO_4 into each grid-box from the synthetic run. Across the whole model ocean interior (all grid boxes below the surface grid-boxes where remineralisation is occurring), the mean proportion of remineralisation flux to the total flux into each grid box is 0.005$\pm$0.025 ($\pm$1 standard deviation). The proportion is generally two orders of magnitude higher at shallower depths (290m) than at depth (2106m) (Figure 4.7a *cf.* Figure 4.7b), reflecting the decrease in remineralisation fluxes with depth whilst circulation fluxes are generally the same magnitude (Figures 4.7c and 4.7d). Even relatively small errors in the circulation flux are therefore likely to dominate over the remineralisation fluxes leading to large
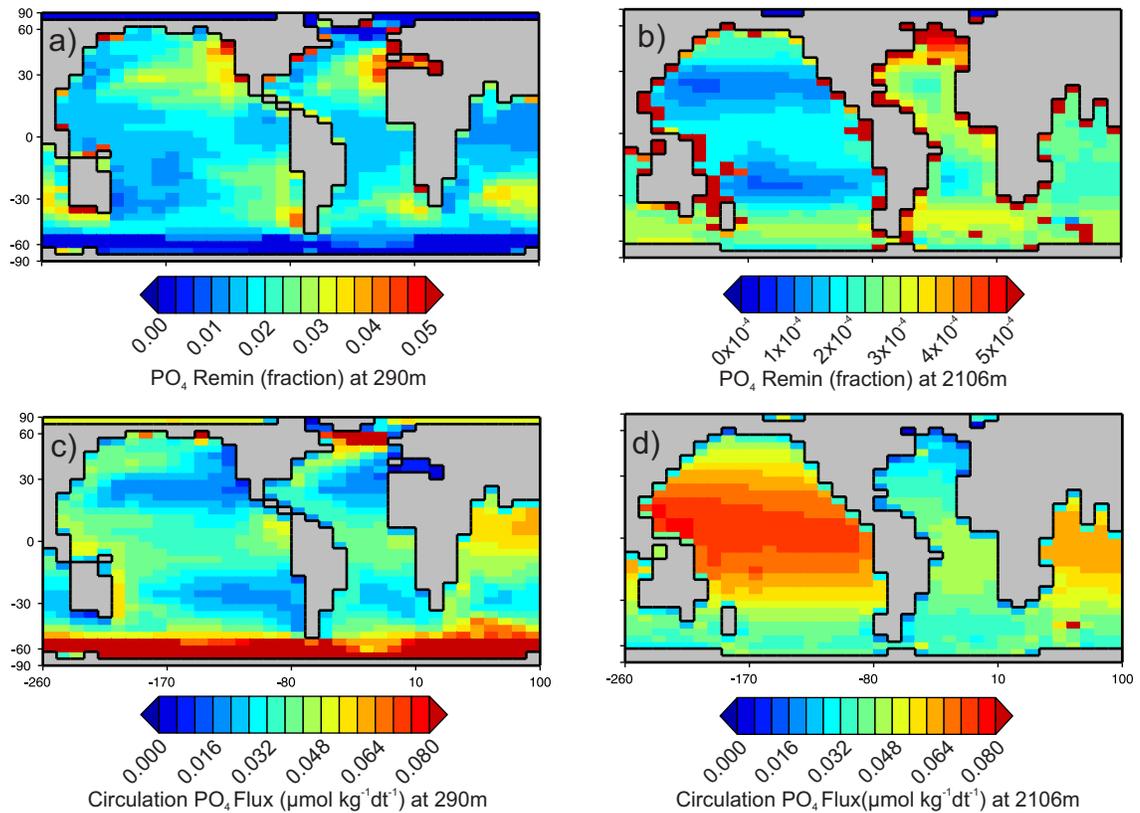
Figure 4.7: Comparison of inputs of PO$_4$ from remineralisation and circulation at steady-state. (a) PO$_4$ remineralisation as a proportion of the total PO$_4$ flux into each grid-box calculated using the synthetic tracer field at 290m and b) 2106m. c) the flux of PO$_4$ into each grid-box from circulation only (μmol kg^{-1} dt^{-1}) from the synthetic tracer fields at 290m and d) 2106m.

errors when using modelled circulation rates.

4.4.2.3   Error Comparison

To compare the magnitude of the various possible errors, the global mean synthetic PO$_4$ remineralisation profile is shown with the global mean and median standard deviations of the ISSs calculated from the ERR-OBS and ERR-CIRC experiments (Figure 4.8). The mean and median standard deviation are both used in this figure because the mean standard deviation of the ERR-CIRC experiments are skewed by the large variability in high latitude regions (Figure 4.6a). Both sources of error are larger at shallower depth and mostly decrease in magnitude with depth although the mean circulatio uncertainty increases below 2000m. The magnitude of uncertainty from the observations is
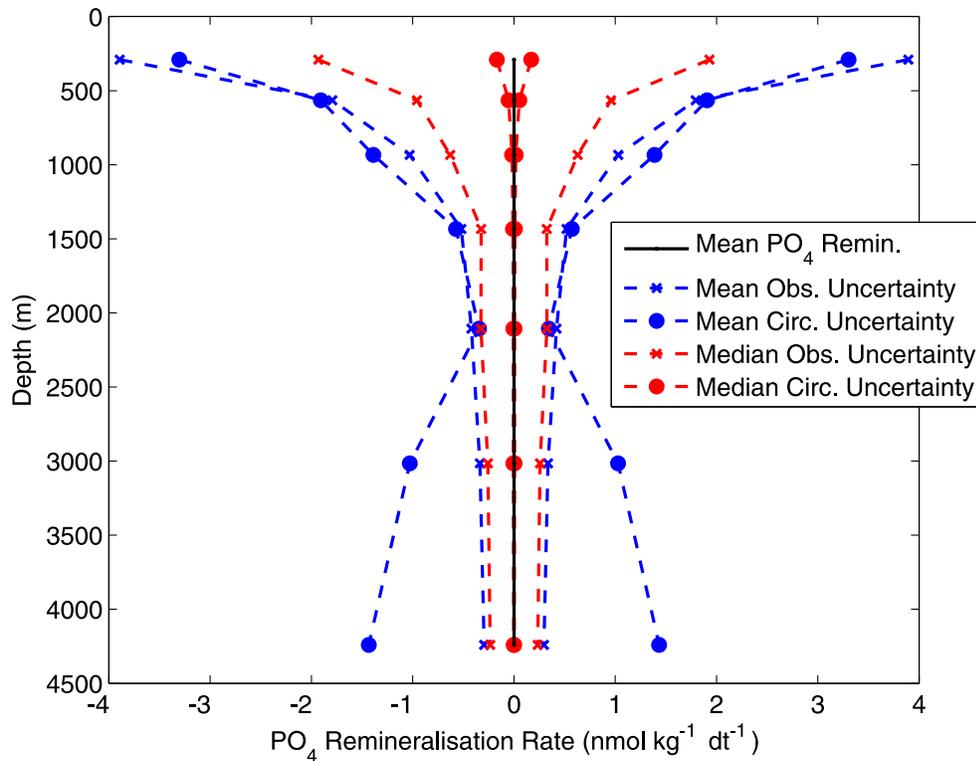
Figure 4.8: Comparison of error magnitudes when estimating remineralisation rates. The global mean PO_4 remineralisation profile from the synthetic dataset is shown with the plus and minus the mean and median standard deviations from the ERR-OBS and ERR-CIRC experiments.

much larger than from the ensemble circulation estimates although the two are similar magnitudes when the mean standard deviation is used to calculate profiles.

Despite similar magnitudes of uncertainty arising from both potential errors in the observations and from the model circulation field, the nature of the uncertainty is different. Uncertainty arising from the observations is partly higher in regions where observations are more uncertain, *e.g.,* coastal areas in Figure 4.5a and 4.5b, but also in regions of the model where convection occurs (Figure 4.5c). In contrast, the uncertainty arising from the model circulation field used may be systematic and dependent on where the model circulation is most different to the real ocean, *e.g.,* Figure 4.6. The patterns in the surface PO_4 ISSs from the MITGCM inversion (Figure 4.2a) show some systematic structure which may suggest that errors are predominantly related to the ocean model rather than from errors in the observations which are likely to be more random. This

is less the case for the deeper ocean (Figure 4.2b).

4.5   Inferring Flux Curves from Remineralisation Rates

4.5.1   Vertical Profiles and Dissolved Organic Matter

In the previous section, a simple approach to estimating remineralisation rates using modelled transport rates has been shown to be sensitive to errors. Taking the next step, in the case that remineralisation rates could be estimated with some reliability, the sensitivity of inferring flux curves by vertically integrating remineralisation rates in the presence of DOM is explored. DOM is used here as an example where the assumption that all remineralisation occurs in the vertical is false. This could also occur from other biological processes such as denitrification. Flux curves are inferred using remineralisation rates from the synthetic dataset (SYN) and a second run where no DOM is exported (SYN-NODOM). To infer a power law curve, a linear trend in fitted to the log transformed remineralisation rates and log transformed depth (*Berelson*, 2001; *Lam et al.*, 2011; *Stanley et al.*, 2012). The gradient of the linear trend gives the value of the exponent for the remineralisation curve, which is converted to a flux curve by adding 1 (*Stanley et al.*, 2012; See also Figure 4.1).

The exponents from power-law curves, fitted to vertical PO_4 remineralisation profiles, when PO_4 is only exported as particulate organic matter, are all close to -1.9 (Figure 4.9a). This corresponds to a flux curve exponent of -0.9, in good correspondence with -0.858 used for the run. The presence of remineralisation from DOM (see Figure 4.3e) lowers the value of the fitted exponent reflecting a shallower Martin Curve (Figure 4.9b). In our experiments, the remineralisation of DOM lowers the fitted exponent by as much as 0.6 to -1.5. This occurs because the remineralisation of DOM inflates the remineralisation in the shallower grid-boxes relative to those in the water column below. In a few grid-boxes in the North Atlantic the estimate of the flux curve expo-

nent increases. This has a strong correspondence with the deepest convection (Figure 4.3f) where DOM is transported deeper into the water column before remineralisation. This highlights that an additional source of remineralisation not restricted to vertical processes can alter the flux curve in an unpredictable way. The bias from DOM occurs predominantly in the high latitudes where DOM is efficiently advected into the ocean interior. The range of exponents purely from this bias is of a similar magnitude and spatial distribution to the Martin Curve exponents found in *Henson et al.* (2012). Other sources of remineralisation may occur in different regions such as in the Arabian Sea, the eastern South Pacific and eastern Tropical North Pacific for dentrification (*DeVries et al.*, 2012b).

4.6   Future Directions

This chapter has presented a straightforward method of using a steady state model circulation, as represented by a transport matrix, to estimate organic matter remineralisation rates from a tracer climatology. The main goal was to explore the feasibility of using this method to infer organic matter flux curves aiding additional understanding of the biological pump in the modern ocean. The results show that this method is associated with a number of significant sources of error that give rise to the spatial patterns and negative values seen in an example inversion using a circulation field from a high resolution ocean model (Figure 4.2). The following sections discuss the potential directions for estimating remineralisation rates from tracer data using model circulation and considerations needed when using these to infer particle flux curves.

The sensitivity to errors in the observations is partly a result of the way that the transport matrix (TM) is constructed. A change in a tracer due to circulation in a model time-step is relatively localised due to the finite speed of advection and diffusion in the model (*Khatiwala et al.*, 2005). Therefore, the uncertainty in the ISS estimates is larger when individual coefficients in the TM are relatively larger (see Table 4.1),
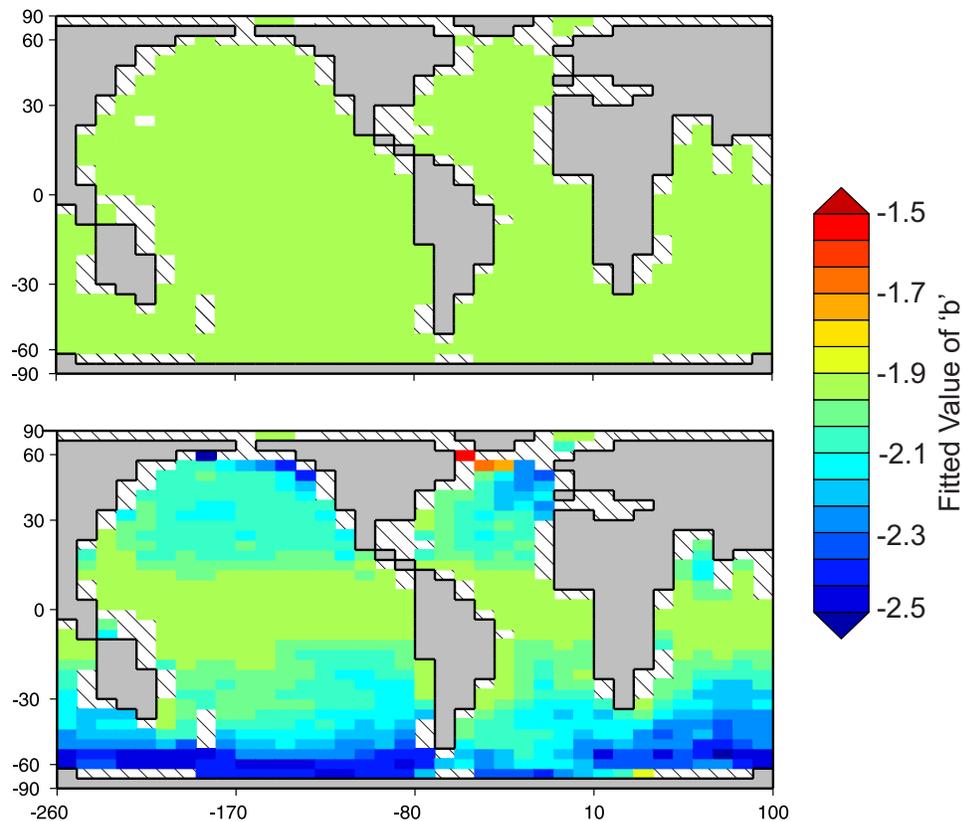
Figure 4.9: Assessment of the uncertainty associated with dissolved organic matter when inferring flux profiles. Value of the exponent when fitting a power law to the water column remineralisation rates from a) a model with remineralisation from only sinking particulate and b) a model with particulate and dissolved organic matter. A value of -0.858 for the Martin Curve was used in both models. All curves were fitted with an $R^2 > 0.9$. The exponent for the remineralisation curve is equivalent to -0.858+1. More negative values indicate a Martin Curve that predicts shallower remineralisation in the water column. Hatched areas indicate where the water column contained too few boxes to fit a remineralisation curve (n<3).

i.e., where there is a larger throughput of water during one timestep, or where there are more grid-box connections such as in regions of convection. This will be a feature of all TMs constructed using the method of *Khatiwala et al.* (2005) regardless of the model used. The ISS estimates are also sensitive to uncertainty in the observations themselves. Previous methods have relied on relating multiple tracers together such that the model transport terms cancel out *e.g., Anderson and Sarmiento* (1994); *Sarmiento et al.* (2002) and is a method which could be applied using the TM. For example, estimating ISSs from two tracers and taking the ratio, *e.g.,* PO_4 and alkalinity to estimate organic carbon to carbonate ratio, may potentially cancel out the systematic

error of the model circulation. The sensitivity to uncertainty in the tracer observations however, may pose a problem for such an approach, especially if observational errors are different between tracers.

The results also illustrate the sensitivity of remineralisation estimates to differences between the model transport rates and actual transport rates in the ocean. Remineralisation fluxes of PO_4 in the synthetic tracer experiments are orders of magnitude smaller than fluxes of PO_4 from circulation. Model transport rates would therefore need to attain a high level of accuracy to minimise the effect of error on the solutions (*Anderson and Sarmiento*, 1994; *Sarmiento et al.*, 2002). Even a data assimilated model, such as produced by the ECCO Consortium ('Estimating the Circulation and Climate of the Ocean'; *Stammer et al.*, 2004), designed to be a dynamically consistent estimate of ocean circulation over recent decades may still lead to large diagnosed flux errors. An additional consideration is whether the model circulation is at steady state. The relative abstraction and speed of GENIE allow an acceptable steady state to be reached which is not always possible in higher resolution and more complex ocean models. This consideration also extends to how representative an annual average TM may be of seasonally varying circulation as opposed to the non-seasonally forced circulation used here.

The flipside of the magnitude of the circulation control on the diagnosed remineralisation rates is that tracers with a steady-state constraint, where it is expected that there should be no significant sources or sinks at depth, could be used to estimate the magnitude of the circulation-based error. As an example, an ISS field generated when inverting the salinity field from the synthetic dataset with the corresponding transport matrix is shown in Figures 4.10a and 4.10b. Salinity ISSs are randomly distributed around zero, consistent with the concept that salinity is not significantly increasing or decreasing in the ocean interior. In comparison, using a different transport matrix, arbitrarily chosen from the ensemble of plausible modern ocean circulations, results in distinct spatial patterns in the ISSs (Figures 4.10c and 4.10d). Comparing the salinity

ISSs (Figure 4.10c and 4.10d) with the errors from the inversion of the synthetic $[PO_4]$ field using the same TM (Figures 4.10e and 4.10f) shows that the patterns in the fields are visibly similar. For example, there are similarities between the positive ISSs for salinity in Figure 4.10d and the negative errors in the PO_4 remineralisation estimates in Figure 4.10f. This suggests that conservative tracers, *i.e.,* tracers that do not have sources or sinks in the ocean interior, could constrain the magnitude of error. For example, considering equation 4.3 but focussing on an individual grid-box, the ISS ($\mathbf{q}_{i=1}$) is a function of the TM coefficients (M_i) and the tracer concentrations (C_i) as per the example in Table 4.1:

$$\mathbf{q}_{i=1} = \sum_{i=1}^{N} M_i C_i \tag{4.5}$$

Redfining the modelled circulation terms to reflect that the modelled circulation is a function of a 'true' circulation term and an error term ($M_i = F_i + \epsilon_i$) and substituting into 4.5 and expanding:

$$\mathbf{q}_{i=1} = \sum_{i=1}^{N} F_i C_i + \sum_{i=1}^{N} \epsilon_i C_i \tag{4.6}$$

For a conservative tracer at steady state, it is expected that $\sum_{i=1}^{N} F_i C_i = 0$. A significant departure from zero in $\mathbf{q}_{i=1}$ is likely to result from the error terms. This may provide a way forward to constrain the ISSs produced by the TM method described here. Such a method would be conceptually similar to the mixing model approaches that use steady state constraints (*Gebbie and Huybers*, 2010). There are however, a limited number of tracers available (*e.g.,* temperature, salinity, $\delta^{18}O$, $\Delta^{14}C$, CFCs) that could realistically be used to constrain model circulation errors. In the GENIE TM, there are typically 15 grid-box connections used for each calculation which would lead to an underdetermined solution, *i.e.,* where the unknowns outnumber the constraints. It would be interesting to see if a simplified TM with fewer coefficients, such as matching the method of *Gebbie*

and Huybers (2010), could use this approach. Ultimately, the number of conservative tracers currently available to perform this calculation is limiting, for example CFCs in the ocean are not in steady-state because of their introduction into the atmosphere as an anthropogenic emission during the twentieth century (*Fine*, 2011). Additionally, other factors would need considering such as whether it is realistic to expect a solution exists (this is explored further in Chapter 5) and whether other constraints, such as remineralisation rates cannot increase with depth, need to be considered.

The focus of estimating remineralisation rates is to use them to infer information about the vertical fluxes of particulate organic matter in different regions of the ocean. Therefore, this chapter has also explored whether, in the event that remineralisation rates could be estimated to some level of accuracy, there are additional sources of uncertainty when inferring flux curves. The presence of additional sources/sinks that are not related to the vertical flux, such as from DOM (section 4.5), could potentially bias the estimation of flux curves. In particular a bias such as DOM, at least in GENIE, causes spatial variability in flux curves that is similar in pattern and magnitude to patterns of particulate organic carbon fluxes found in sediment trap studies (*Henson et al.*, 2012; *Marsay et al.*, 2015). Any method of estimating flux curves, whether derived purely from observations (*Stanley et al.*, 2012) or using modelling circulation rates, need to consider this issue. The relatively low vertical resolution of GENIE (8 vertical levels: Chapter 2) may exaggerate this problem because grid-boxes with additional sources/sinks are weighted more heavily in the calculation. However, this is still an important issue to consider for future approaches.

An alternative approach to exploring the spatial variability in particulate organic matter fluxes using modelled circulation is the optimisation of parameters in forward models, *e.g.,* (*Schlitzer*, 2004; *Kwon and Primeau*, 2006; *Kriest et al.*, 2012; *Teng et al.*, 2014). Such an approach can explicitly include DOM remneralisation, circumvent the problems of directly inverting a tracer described in this chapter, and circumvent the need to infer flux curves from vertical profiles of remineralisation rates. Transport matrices are a
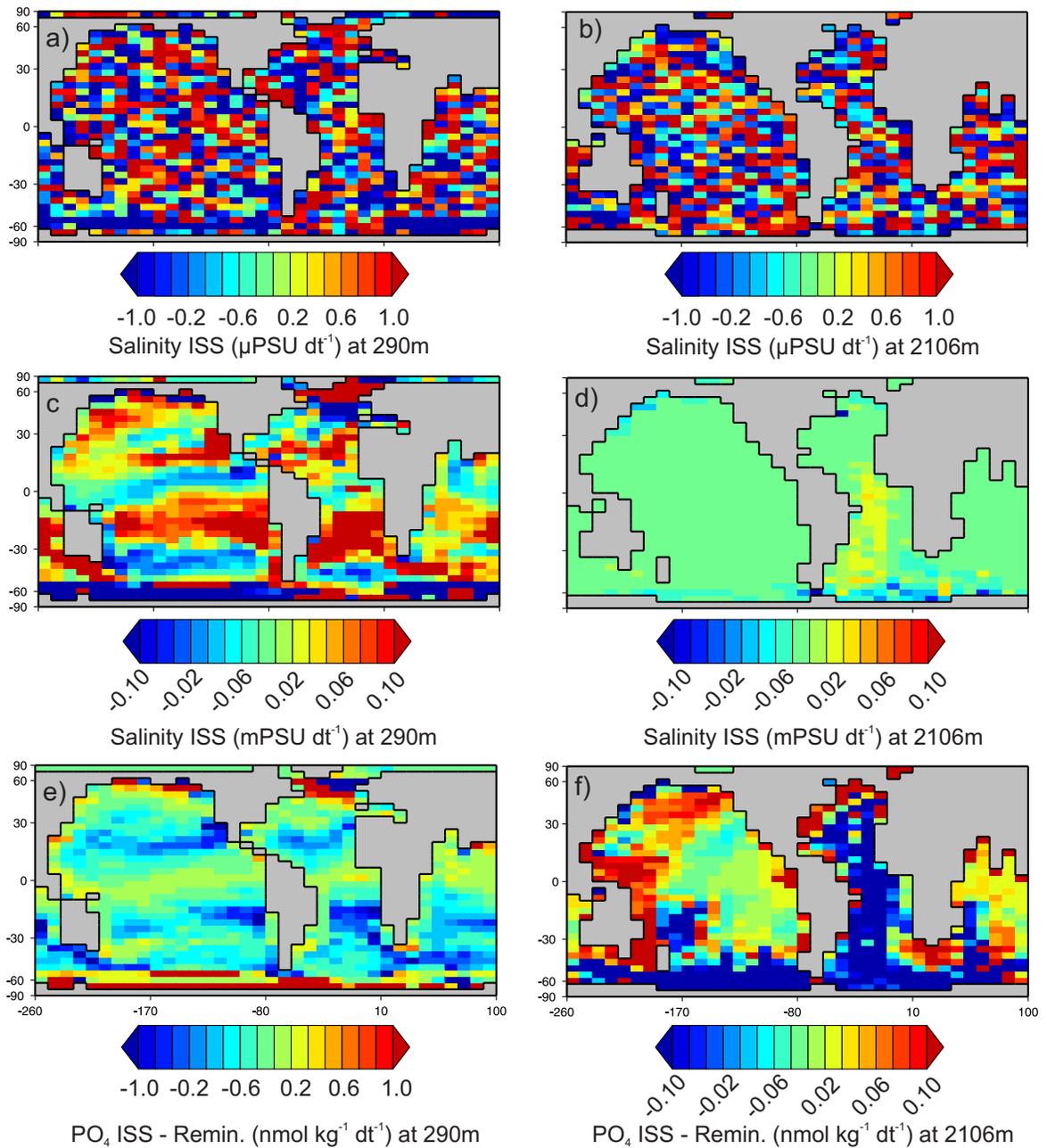
Figure 4.10: Inversion of salinity as a possible constraint on the uncertainty from using a modelled circulation. (a) Inversion of the salinity field from the synthetic dataset using the corresponding transport matrix at 290m (PSU dt^{-1}) and at (b) 2106m. (c) Inversion of the salinity field using a alternative transport matrix from the ensemble of plausible modern ocean circulations (PSU dt^{-1}) at 290m and (d) 2106m. (e) The error of the synthetic PO$_4$ ISS (mol kg^{-1} dt^{-1}) (ISS-synthetic) using the same transport matrix in panels c and d at 290m and f) 2106m.

useful tool for this approach to help avoid the length multi-thousand year integration times for a tracer field to reach equilibrium. In the next chapter, such an approach based around the GENIE transport matrix described in this chapter, is described and applied to observed [PO$_4$].

CHAPTER 5

Optimisation of regionally variable organic matter remineralisation depths in an Earth System Model

5.1 Introduction

Sinking particles of organic matter transfer 5-10 Pg C per year of carbon (and associated nutrients) from the upper ocean to the ocean interior, in a process known as the biological pump (*Henson et al.*, 2012). As these particles sink, they are remineralised through bacterial and zooplankton related processes, releasing the carbon and nutrients back into solution at depth. Vertical fluxes of particles in the water column have been historically described by the canonical Martin Curve, a power-law function that describes a rapid decrease in flux at shallow depths leading to a small asymptotic flux at deeper depths (*Martin et al.*, 1987). The power-law exponent (b) controls whether organic matter is remineralised predominantly at shallower or deeper depths. The e-folding depth has been previously used to describe differences in the functions that model vertical organic matter fluxes (*Kwon et al.*, 2009). This is the depth at which the exponentially decreasing flux has been reduced to $\frac{1}{e}$; approximately 0.37 times the initial flux exported from the surface (conversely this is the depth at which $\sim$63% of the flux has been remineralised). In this chapter, the term 'remineralisation

depth' refers to the e-folding depth of the function used to represent the vertical flux of particulate organic matter. The Martin Curve has an exponent value of b=-0.858, corresponding to an e-folding depth of 319m. The Martin Curve, and other functions such as exponential functions: see dashed profiles in Figure 5.1, are commonly used by ocean biogeochemical models as way of parameterising water column remineralisation of sinking particles. Traditionally this has been applied with a single global exponent such that the depth of remineralisation is treated as a global average.

Recent analyses of global sediment trap observations have started to suggest a spatially heterogeneous pattern in Martin Curves that varies predominantly by latitude (*Henson et al.*, 2012). Particulate organic matter (POM) fluxes in high latitudes decrease faster with depth (*i.e.,* shallow remineralisation functions: b=-1.18) than in low latitudes where a greater proportion of POM is delivered to depth (b=-0.24) (Figure 5.1). A number of hypothetical mechanisms have been proposed to explain these spatial patterns, including the increased sinking velocity of particles via a 'ballast effect' (*Armstrong et al.*, 2002; *Klaas and Archer*, 2002), aggregation of particles (*Passow*, 2004) and surface ecosystem processes such as the level of ecosystem activity and how labile organic matter is (*Lam et al.*, 2011; *Le Moigne et al.*, 2012). However, further progress on a mechanistic understanding has been hindered by the relatively low sampling density of sediment traps in space (and time), particularly in the Southern Ocean (*Le Moigne et al.*, 2012) (see Chapter 3).

Ocean biogeochemical models predict that atmospheric CO_2 is sensitive to changes in the global remineralisation depth. This can occur via the redistribution of dissolved organic carbon (*Kwon et al.*, 2009) and changes in the whole ocean carbon chemistry via changes in the calcium carbonate ($CaCO_3$) content of deep-ocean sediments over longer time-scales (*Archer and Maier-Reimer*, 1994; *Roth et al.*, 2014). These processes could potentially respond to a range of different environmental parameters that could lead to feedbacks of different magnitude and directions. The potential impact of remineralisation depth changes on atmospheric CO_2 and climate is a key uncertainty for
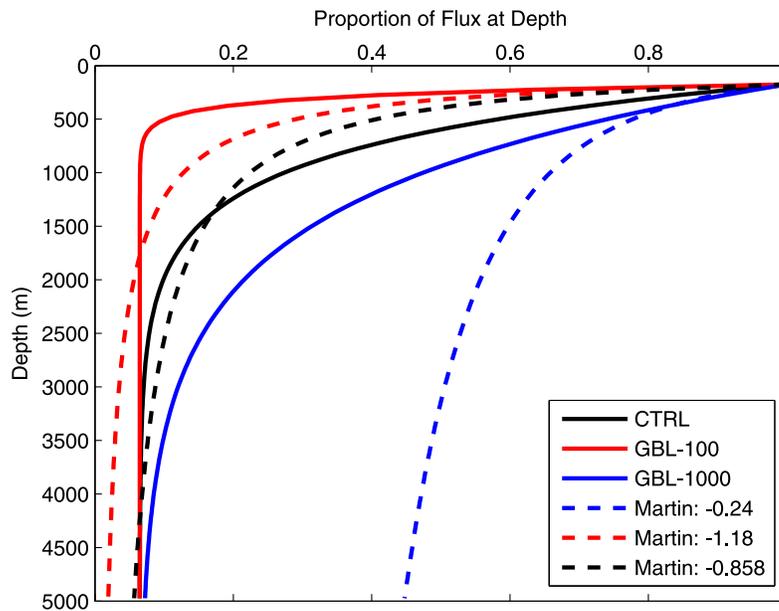
Figure 5.1: The normalised water column distribution of particulate fluxes defined using the exponential functions used in this study (equation 5.5). As a comparison, Martin Curves are shown with the minimum and maximum exponent values found when fitting curves to sediment trap data by *Henson et al.* (2012). All curves are normalised to an export depth of 175m consistent with the bottom depth of the surface grid-box in the 8 vertical level configutation of GENIE used here.

understanding both past and current changes in climate (*Honjo et al.*, 2014).

Modelling studies varying key global biogeochemical parameters, such as the rate of nutrient uptake and remineralisation depth, have shown that [PO$_4$] distributions are particularly sensitive to changes in the global mean remineralisation depth (*Schlitzer*, 2002a, 2004; *Kwon et al.*, 2009; *Kriest et al.*, 2010, 2012). An optimal value for the remineralisation depth can therefore be found by statistically comparing the model [PO$_4$] field to observed [PO$_4$]. These studies found that the Martin Curve exponent is close to -0.9, similar to the value found by (*Martin et al.*, 1987), although lower than the mean value of -0.639 found by (*Henson et al.*, 2012). [PO$_4$] distributions are also more sensitive to global changes in remineralisation depths than to export production and dissolved organic matter production (*Kwon and Primeau*, 2006; *Kriest et al.*, 2012). This approach of optimising biogeochemical parameters is an important method for quantifying the biological pump that can complement the observations from sediment trap data.

Parameter optimisation studies have previously accommodated spatial variability in remineralisation depths. *Schlitzer* (2002a) used an adjoint method to systematically vary ocean circulation, air-sea fluxes, production and remineralisation rates simultaneously in the Southern Ocean (<50°S) to determine parameter values that were in good agreement with observations. The adjoint method calculates the gradient of a statistical fit to observations with respect to the parameters that can be used to generate a new set of parameter values by minimising the gradient (*Schlitzer*, 2002a, 2004; *Yao and Schlitzer*, 2013). A single Martin Curve was applied to the whole Southern Ocean with b as a tunable parameter. Average values of b=-1.04±0.10 were found across a range of experiments, providing a quantitative test of the sensitivity to model structure (*Schlitzer*, 2002a). This value is in agreement with values in the Southern Ocean found by *Henson et al.* (2012). The same method was also applied to the Equatorial and North Pacific in *Schlitzer* (2004) and globally in *Yao and Schlitzer* (2013), but with remineralisation depths that are able to change in each water-column independently rather than as a regional mean value. Unfortunately the results for remineralisation depths were not reported (*Schlitzer*, 2004; *Yao and Schlitzer*, 2013). A similar adjoint method analysis was used by *Usbeck* (1999) for the Atlantic Ocean. The estimated remineralisation depths were similarly shallow in the Southern Ocean, agreeing with *Schlitzer* (2002a) and *Henson et al.* (2012). However, they were also shallow in the Equatorial Atlantic and deeper in the North Atlantic, opposite to the latitudinal patterns reported by *Henson et al.* (2012) (Figures 5.2d *cf.* 5.2b).

Nutrient distributions have also used to evaluate and assess the incorporation of mechanistic parameterisations of remineralisation in models, *e.g., Howard et al.* (2006) and *Chikamoto et al.* (2012). *DeVries et al.* (2014) developed a mechanistic model of particle sinking based on particle size distributions that was evaluated against observed [PO_4]. Again, there are notable differences with *Henson et al.* (2012), with greater particle fluxes reaching the deep ocean in the high latitudes (Figures 5.2c *cf.* 5.2a). This also partly disagrees with the findings of *Usbeck* (1999) where North Atlantic is a region
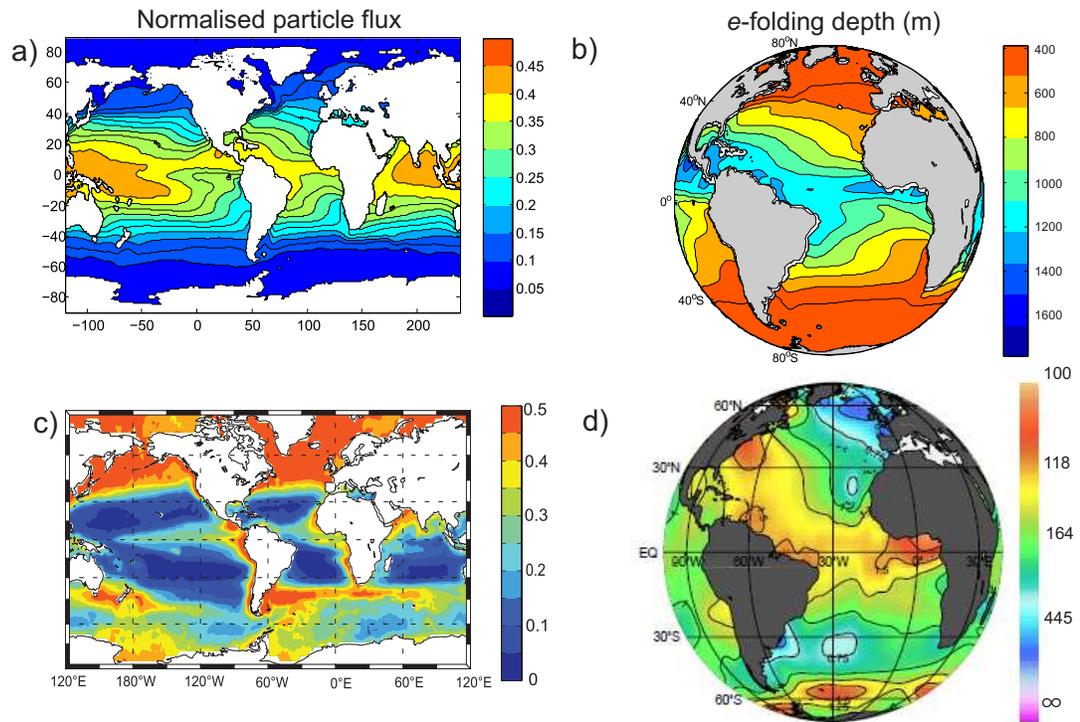
Figure 5.2: Comparison of remineralisation depth metrics from a recent sediment trap analysis and two model optimisation studies. (a) normalised particle flux at 2000m from sediment trap analysis adapted from *Henson et al.* (2012), (b) *e*-folding depths from Martin Curves fitted to sediment trap data adapted from *Henson et al.* (2012). (c) normalised particle flux at 1000m adapted from the mechanistic model of *De Vries et al.* (2014), (d) *e*-folding depths of power law functions adapted from the optimisation study of *Usbeck* (1999). Panels a and c are comparable but note the differences in depth. Flux curves are almost linear at depth (*Honjo et al.*, 2008) such that the spatial patterns will not differ greatly. Panels b and d are comparable.

of shallow remineralisation. These two studies are not strictly comparable as multiple parameters, (*e.g.,* mixing coefficients, surface heat fluxes, gas exchange rates as well as organic matter export and remineralisation depth) and tracers, (*e.g.,* PO_4, alkalinity and oxygen) were used in *Usbeck* (1999), but it seems reasonable to expect similar results given that the global remineralisation depth is a key parameter controlling nutrient distributions and similarities between large-scale circulation patterns. Ultimately, the results of the optimisation and modelling studies described do not agree with the findings from *Henson et al.* (2012) and provide a wide range of solutions (Figure 5.2).

5.1.1   Aims

In this chapter the spatial variability in remineralisation depths is explored in the context of a model optimisation study. The aim of the chapter is to explore whether there is an optimum fit to observed [PO$_4$] in a similar manner to previous studies (*Kwon and Primeau*, 2006; *Kriest et al.*, 2012) that can be used to inform mechanistic interpretations along with sediment trap studies. The chapter also has an additional focus on assessing how well the remineralisation depths are constrained in light of the range of previous results in Figure 5.2.

To achieve this, a new steady-state model of phosphorus cycling based on GENIE is built around the GENIE transport matrix developed in the previous chapter. A set of 15 oceanic regions is defined, within which the remineralisation depth can vary independently of other regions during a model run. A Latin hypercube is used to efficiently sample across the large number of potential combinations of spatially variable remineralisation depths that are then used to simulate [PO$_4$] fields in the new model. By covering the range of all possible solutions, the Latin hypercube approach can be used to explore whether there are multiple solutions that can produce model [PO$_4$] fields that are equal fits to observed [PO$_4$] as well as assessing how the remineralisation depth in each region contributes to the fit to observations.

5.2   Model and Experiment Description

5.2.1   Model Description

The number of model runs (individual sets of parameters) required to cover a multi-dimensional parameter space increases rapidly with the number of parameters and the resolution of the sampling, *i.e.,* 'the curse of dimensionality'. The Latin hypercube approach, described further in section 5.2.2, is a method of random sampling that effi-

ciently fills the parameter space for a given number of samples. Even with this method, multiple samples are still needed to ensure that the parameter space is representatively sampled, requiring a large number of model runs. To make this feasible, a new steady-state phosphorus-only model based on the 8-level model configuration of the Earth System model GENIE (*Ridgwell et al.*, 2007a: Chapter 2) is first developed. This model is similar to previous models used by *Kwon and Primeau* (2006) and *DeVries et al.* (2014). Although GENIE is itself, relatively fast, the aim is to construct a significantly faster 'offline' version that removes significant computing constraints associated with the Latin Hypercube approach used here. A full 10,000 year spin-up of GENIE is first performed to reach a steady-state solution for the pre-industrial ocean, taking approximately 15 hours. This run, hereafter referred to as the control run (CTRL: Table 5.2), forms the basis for constructing the offline model described in the following sections.

5.2.1.1   Circulation Model

A transport matrix (**A**) is diagnosed from the CTRL circulation at the end of the 10,000 year spin-up using the method based on *Khatiwala et al.* (2005) described in Chapter 4 and Appendix B. The circulation in the 8-level version of GENIE used here is non-seasonally forced. The transport matrix represents the change in a tracer due to the effect of advection, diffusion and convection calculated within a single time-step of GENIE. In contrast to its application in Chapter 4, here the transport matrix is used as the steady-state ocean circulation for a biogeochemical model.

The transport matrix diagnosed from GENIE includes the effect of virtual salinity fluxes and so the biogeochemical tracers are normalised by salinity before multiplication with the transport matrix and converted back to concentration units afterwards. This means that the mass of tracers is not exactly conserved requiring a rescaling of the tracer concentrations to ensure there has been no net gain or loss, see *Edwards and Marsh*

(2005) and *Ridgwell et al.* (2007a).

5.2.1.2   *Biogeochemical Model*

The model contains two dissolved tracers, PO_4 and dissolved organic phosphorous (DOP) which are governed by the following equations:

$$PO_4^{t+1} = \mathbf{A}PO_4^t - Juptake_{z<z_0} + JPOP_{z>z_0} + JDOP \qquad (5.1)$$

$$DOP^{t+1} = \mathbf{A}DOP^t + (v \cdot Juptake_{z<z_0}) - JDOP \qquad (5.2)$$

where t and $t+1$ refer to the timestep (as per equation 4.2) and z refers to depth and z_0 to the bottom of the surface grid-box. The uptake of dissolved PO_4 in the surface by phytoplankton (*Juptake*) is calculated using the scheme described by *Ridgwell et al.* (2007a) (see Table 5.1) where a maximum uptake rate of PO_4 ($u_0^{PO_4}$) is kinetically limited by a Michaelis-Menton type limitation using a half-saturation constant (K^{PO_4}). Limitation from other nutrients, such as Iron which is important in the Southern Ocean, is not included in this model. Uptake is also modified by sea-ice cover (F_{seaice}) and the strength of the local insolation (I) which is normalised by the solar constant (I_0). Sea-ice cover and insolation are taken from the CTRL run and are fixed in the model:

$$Juptake = u_0^{PO_4} \cdot \frac{PO_4}{PO_4 + K^{PO_4}} \cdot (1 - F_{seaice}) \cdot \frac{I}{I_0} \qquad (5.3)$$

A fraction of the PO_4 uptake (v) is partitioned into DOP which can be transported by ocean circulation and is remineralised with a time constant of ($1/k$) (Table 5.1):

$$JDOP = kDOP \qquad (5.4)$$

The remaining fraction of uptake $(1 - v)$ is exported vertically as sinking particles of organic matter that are partitioned into a labile fraction $(1 - f^{POP})$ and a refractory fraction (f^{POP}). The refractory fraction is analogous to the 'protected' fraction of organic matter in the ballasting hypothesis (equation 1.9). The two fractions has been used previously and found to produce a better fit to tracer observations (*Andersson et al.*, 2004; *Ridgwell et al.*, 2007a; *DeVries et al.*, 2014). The remineralisation of POP is instantaneous, *i.e.,* it occurs within the same time-step as the export flux. The flux of POP left at depth (z) proceeds according to an exponential decay:

$$\text{JPOP}(z) = f^{POP} \cdot \text{JPOP}(z_0) \cdot exp(\frac{z_0 - z}{l^{rPOP}})$$
$$+ (1 - f^{POP}) \cdot \text{JPOP}(z_0) \cdot exp(\frac{z_0 - z}{l^{POP}}) \quad (5.5)$$

The attenuation of POP with depth is controlled by l^{POP}. The remineralisation of a smaller fraction of refractory POP (0.065 of the exported flux from the surface) is calculated using a longer length scale (Table 5.1) that results in the majority of the refractory POP reaching the deep ocean. There is no sediment component in this model so POP that has reached the sediment is remineralised fully in the grid-box directly above. This maintains a closed-system in respect to PO_4. An alternative approach, as described by *DeVries et al.* (2014), is to remove all POP reaching the sediment and matching it with an allochthonous source to the ocean. This would have the effect of reducing the total remineralisation in the deep ocean and potentially lead to different results. The approach in GENIE is retained here to maintain direct comparability.

A simulation is achieved by time-stepping the equations forward in time. The ocean circulation is calculated using a time step of 0.01 year whilst the biogeochemistry is calculated once every 5 time-steps of the ocean model (0.05 year), as per *Ridgwell et al.* (2007a). Previous studies have alternatively used a Newton-Krylov method to

Table 5.1: Values of biogeochemical parameters used in the steady-state model. [a] indicates the value used in *Ridgwell and Hargreaves* (2007) which is altered in this study. Numbers in square brackets refer to the corresponding references.

Parameter	Value	Description
$u_0^{PO_4}$	1.96 μmol kg^{-1} yr^{-1}	Maximum PO$_4$ uptake rate [1]
K^{PO_4}	0.22 μmol kg^{-1} yr^{-1}	Michaelis-Menton half-saturation concentration [1]
k	0.5 yr^{-1}	DOM remineralisation rate [2][3]
v	0.66	Fraction of DOM from PO$_4$ uptake [1]
f^{POP}	0.065	Initial fraction of refractory POM [1]
l^{POP}	550[a] m	e-folding depth of labile POM remineralisation [1]
l^{rPOP}	100000 m	e-folding depth of refractory POM remineralisation [1]
z_0	175 m	Bottom of surface grid-box [2]

[1] *Ridgwell and Hargreaves* (2007) [2] *Ridgwell et al.* (2007a) [3] *Najjar et al.* (2007)

find equilibrium solutions to the governing equations, *i.e.,* finding a solution where ($\frac{dPO_4}{dt} = 0$: *Kwon and Primeau*, 2006; *DeVries et al.*, 2014). A time-stepping method is used here to achieve a steady-state [PO$_4$] distribution in order to maintain consistency with the full GENIE model. A comparison of corresponding runs in both models shows that the steady-state model is a good approximation (r>0.99) for the full GENIE model (Figure 5.3). There are minor differences in the DOM fields (Figure 5.3c and 5.3d). In both figures the negative concentrations of DOM can be found in GENIE but not in the steady-state model. This occurs in the Arctic Ocean and is a consequence of differences in details of the numerical implementation between each model. Other deviations from the 1:1 line in Figures 5.3c and 5.3d occur at shallow depths in the Southern Ocean concurrent with convection in the model. As suggested in Chapter 4, this may reflect minor differences in convection between timesteps that have been sampled by the sequential grid-box dying method used here. The steady-state model is run in MATLAB 2013a, and takes approximately 2 minutes to simulate 3000 years on a standard PC workstation, compared to ~4.5 hours in GENIE. Experiments are run in parallel using the parallel toolbox in MATLAB reducing the total simulation time for multiple experiments.
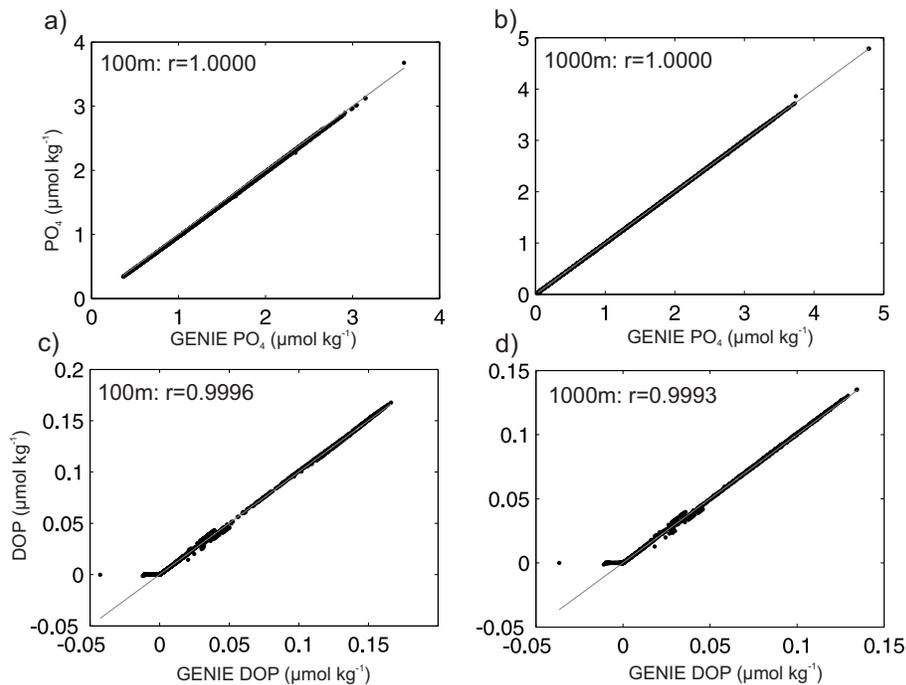
Figure 5.3: Comparison of the [PO$_4$] and DOP fields after a 3000 year simulation using the full GENIE model and the steady-state model described in this chapter. Panels (a) and (c) show [PO$_4$] and DOP when the global remineralisation depth has been changed to 100m from the global value of 550.195m. Panels (b) and (d) show the same but when remineralisation depths have been changed to 1000m.

5.2.2 Experiment Design

5.2.2.1 Defining Regions

A set of oceanic regions are defined to approximately encapsulate the large-scale variability in biogeochemistry and patterns of remineralisation depths observed in sediment trap studies and in Chapter 3. Previous studies have defined model regions for similar purposes based on variables such as vertical mixing, mixed layer depths, sea ice and sea surface temperature (*Sarmiento et al.*, 2004b; *Henson et al.*, 2010; *Fay and McKinley*, 2013, 2014). As ocean circulation is responsible for delivering nutrients to the surface, helping to drive major feature of dissolved nutrients, this approach produces biogeochemical regions that are internally consistent with the model. However, GENIE does not simulate mixed layer depths making this approach difficult to implement (note that more recent developments incorporate the effects of a diagnosed mixed layer depth,

e.g., Matsumoto et al., 2008). As an alternative, regions are defined by lines of latitude and basins similar to the approach used by air-sea flux inversion studies, *e.g., Gloor et al.* (2001) and *Mikaloff Fletcher et al.* (2006). 15 regions are defined based on a partitioning by *Gloor et al.* (2001) with some minor changes (Figure 5.4a). The Arctic Ocean is excluded as a region as little export production occurs due to the extent of sea ice in the model. The Mediterranean is also excluded as it is a relatively shallow (mostly <1000m) and contained basin where remineralisation depth changes will have a minimal effect on the global distribution of $[PO_4]$. The defined regions broadly correspond with the major features in the surface $[PO_4]$ field from the CTRL run such as higher concentrations in upwelling regions and lower concentrations in the nutrient-depleted gyres (Figure 5.4b). This suggests the regions should be a reasonable analogue for the approach using circulation metrics. Regions are also comparable to the ocean biomes used in the TransCom: the Atmosphere Tracer Transport Model Intercomparison Project, *e.g., Gurney et al.* (2003). Alternative choices of regions were also tested including latitude bands that were not separated by basin (n=7) and higher resolution regions (n=29) (see Appendix C: Figures C.7 and C.8).

5.2.2.2   Model Experiments

The remineralisation length scale (l^{POP}) is allowed to vary individually between 100m and 1000m in each of the 15 regions (compared to 550m in the control run), illustrated by the solid profiles in Figure 5.1. The range for l^{POP} used here is similar to ranges used in previous studies using GENIE (*Ridgwell et al.*, 2007a; *Holden et al.*, 2013). This range reflects a range of *e*-folding depths of 275m to 1170m, comparable to the range found by *Henson et al.* (2012) of 232m to 6297m (calculated from the range of Martin Curve *b* values of -1.18 to -0.24). The deeper *e*-folding depths from the estimates of *Henson et al.* (2012) reflect a potentially anomalously deep end-member estimate as also found in its use to calculate synthetic fluxes in Chapter 3 (see discussion of Figure 3.12). The appropriateness of the range used is explored in the results. A sensitivity
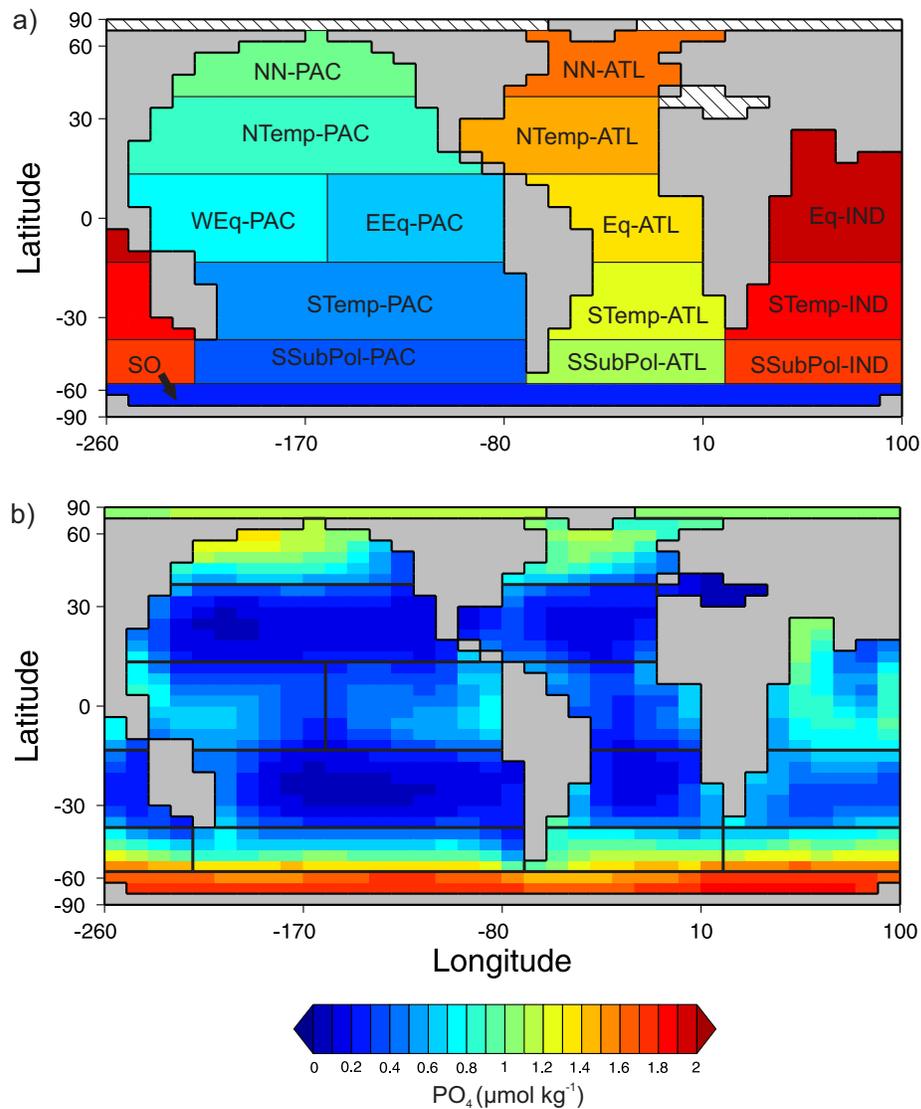
Figure 5.4: The 15 regions used for the Latin hypercube sampling. (a) Location and names of the 15 regions on the GENIE grid, based on regions from *Gloor et al.* (2001). Boundaries are at 58°S, 36°S, 13°S, 13°N, 36°N. The equatorial Pacific is split at 198.75°E following *Mikaloff Fletcher et al.* (2006). The remineralisation depth in each region is allowed to change independently of the others. The Arctic and Mediterranean, shown in hatching, are excluded and take the global average value of 550.195 in all experiments (Table 5.1). (b) Surface [PO$_4$] field from the control run of GENIE (CTRL; Table 5.2) with the region boundaries superimposed.

test on the fraction of refractory POM (f^{POP}: equation 5.5) showed that [PO$_4$] was much less sensitive to changes in this parameter compared with remineralisation depths (Figures C.5 and C.6).

A sensitivity analysis is first performed to explore the behaviour of [PO$_4$] in response to perturbations to the length scale in each region individually. The length scale is assigned to both 100m (SENS-100) and 1000m (SENS-1000) in each region separately whilst the length scale in the other regions are kept constant at 550m (Table 5.2). Experiments where the length scale is set to 100m and 1000m globally are also performed, *i.e.*, in all regions, as a comparison (GBL-100 and GBL-1000: Table 5.2).

Latin hypercube sampling is used to vary the length scales in all regions simultaneously with the aim to sample across the range of all possible combinations of spatially variable remineralisation depths such that optimum combinations can be identified. Random and grid (or systematic) sampling are alternative methods to sample model output over a range of input parameter values. However, multiple sets of random samples cannot guarantee that the parameter space is sampled evenly. For example, Figure 5.5a shows a scenario where random sampling can produce samples in close vicinity to each other resulting in an inefficient and unrepresentative sample of all possible combinations. Grid sampling, whilst covering the parameter space evenly (Figure 5.5b), requires an increasingly large number of samples with every parameter added and increase in sampling resolution, *e.g.*, for just 4 samples of each of the 15 remineralisation depths, this would require 50625 runs, taking $\sim$ 70 days to run with the steady-state model: the 'curse of dimensionality' (*Urban and Fricker*, 2010).

In a Latin hypercube, the parameter range is partitioned into equally spaced ranges equal to the size of the number of samples required. Samples are chosen randomly with the constraint that each row and column can contain only one sample (comparable to the principle of a sudoku puzzle) (Figure 5.5c). This quality of Latin hypercubes ensures a more efficient sampling across the parameter space than the random sampling on its
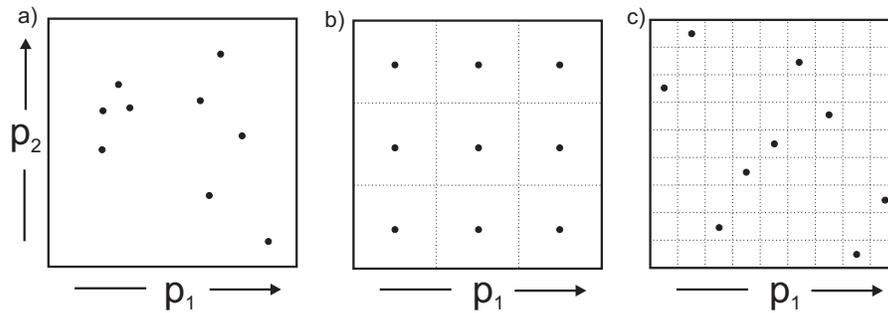
Figure 5.5: Schematic of parameter sampling methods for two parameters (p_1 and p_2) for n=9 samples. (a) random sampling, (b) gridded sampling and (c) Latin hypercube sampling. Adapted from *Urban and Fricker* (2010).

own for any given number of samples. The number of samples is defined by the user. The key advantage of the Latin hypercube is that the sampling scales with computing resources (*Urban and Fricker*, 2010), *i.e.,* the parameter space is always filled efficiently when using smaller smaples but it can be sampled with increasing resolution with the ability to take more samples. Due to its relatively fast speed of integration, the steady-state model described in this chapter can therefore take full advantage of the Latin hypercube sampling by sampling the parameter space with more resolution providing an approach not available from the 'online' GENIE model or higher resolution models.

The length scales are sampled from a uniform prior distribution ranging from 100m to 1000m using the *lhsdesign* function in MATLAB 2013a, *i.e.,* the Latin hypercube partitioning is equal such that there is an equal probability of the length scale being between 100 and 1000m (see Figure C.4). A general rule-of-thumb states that the number of samples should be at least 10 times the number of parameters (*Loeppky et al.,* 2009; *Gregoire et al.,* 2011; *Williams et al.,* 2013). The steady-state model permits an increase in the number of samples to 40 times the number of parameters, resulting in a total of 600 simulations. More simulations could be performed but introduces problems with storing and analysing a large dataset. All experiments are run for 3000 years from the CTRL experiment (see Table 5.2). Annual mean fields of [PO_4] are diagnosed at 2999.5 years. A simulation of 3000 years is sufficient for the mean surface and deep ocean [PO_4] to equilibrate to a global change in the remineralisation length scale to

Table 5.2: Description and names of the experiments used in this chapter.

Experiment	Description
CTRL	Control experiment using parameter values from Table 5.1
GBL-100	Experiment where $l^{POP} = 100$m globally
GBL-1000	Experiment where $l^{POP} = 1000$m globally
SENS-100	Sensitivity experiments - $l^{POP} = 100$m for each region individually
SENS-1000	Sensitivity experiments - $l^{POP} = 1000$m for each region individually
LHS	Latin hypercube ensemble - l^{POP} varies between 100m and 1000m

either 100 or 1000m (see Figure C.3).

5.2.2.3   [PO$_4$] Observations and Cost Function

The model [PO$_4$] distribution is compared to observed [PO$_4$] concentrations from the World Ocean Atlas 2009 (*Garcia et al.*, 2010), regridded onto the GENIE grid using the box averaging function in Ocean Data View 4.4.4 (*Schlitzer*, 2002b). The model-observation misfit is quantified using the global root mean square difference (RMSD):

$$RMSD = \sqrt{\frac{\sum (M - O)^2}{N}} \qquad (5.6)$$

The RMSD is also used to quantify how different sensitivity experiments are to the control. In this case, the RMSD is calculated with the control field in place of the observations.

5.3   Sensitivity of Model [PO$_4$] to Regional Changes in Remineralisation Depth

The sensitivity of the global distribution of [PO$_4$] to regional changes in the remineralisation depth is tested by perturbing the remineralisation depth to the minimum and maximum values in each region individually whilst the remineralisation depth is main-

tained at 550m in all other regions (SENS-100 and SENS-1000 experiments). Basin zonal averages of [PO$_4$] are shown at shallow (290m) and deep (2106m) as a proportional deviation from the CTRL zonal average (Figures 5.6 and 5.7). As a comparison, the zonal averages from the GBL-100 and GBL-1000 runs are also shown. Zero indicates the zonal mean is the same as the CTRL zonal average whilst a deviation of 1 indicates the zonal mean is 100% larger. The model [PO$_4$] distribution is relatively insensitive to regional changes in remineralisation depth in comparison to global changes in remineralisation depth (black dashed lines) irrespective of whether remineralisation is shallower or deeper. Across all the experiments, the largest deviations from the CTRL zonal average in each basin occur when remineralisation changes are local to that basin (Figure 5.6 and 5.7). The largest deviations in each basin occur in a similar latitudinal range as the regional change in remineralisation. For example, the pattern of largest deviations in zonal averages at 290m in the Atlantic follows changes in each region (red dashed, green dashed and yellow dashed lines in Figure 5.6b). Differences in [PO$_4$] occurring in a basin other than that in which the remineralisation depth is changed are difficult to distinguish from the control run zonal averages and are difficult to distinguish even when using a log scale (not shown).
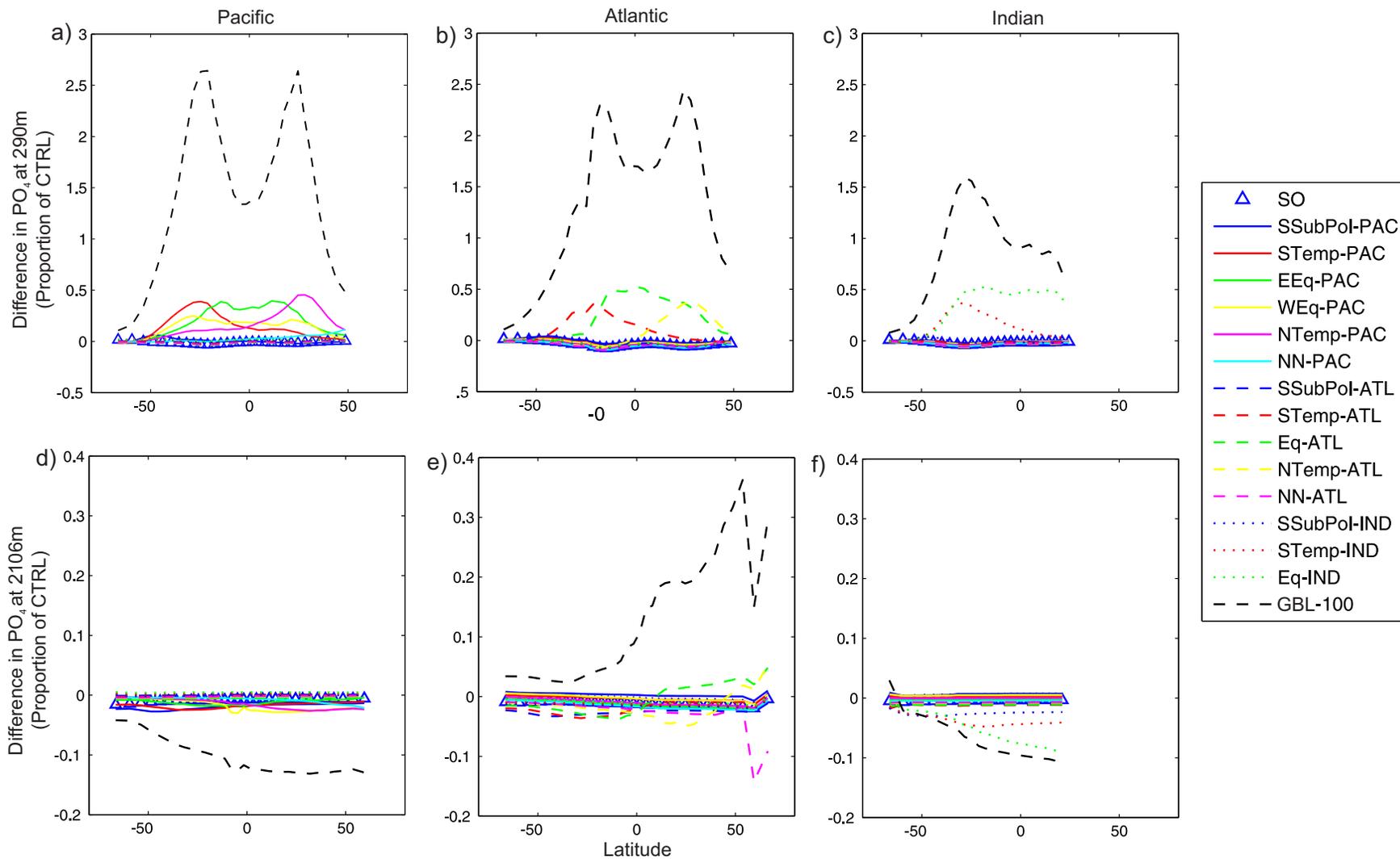
Figure 5.6: Proportional deviation zonal averages of [PO$_4$] from the sensitivity experiments where the remineralisation depth in each region is set to 100m individually (SENS-100; Table 5.2). Each line indicates the proportional change in the zonal average from the CTRL due to a change in the region indicated in the legend. The black dashed line indicates the run where remineralisation depths in all regions are at 100m (GBL-100). The top 3 panels (a,b,c) are zonal averages at 290m, and the bottom 3 panels (d,e,f) are at 2106m. The left two panels (a & d) are for the Pacific, (b & e) for the Atlantic, and (c & f) for the Indian Ocean.
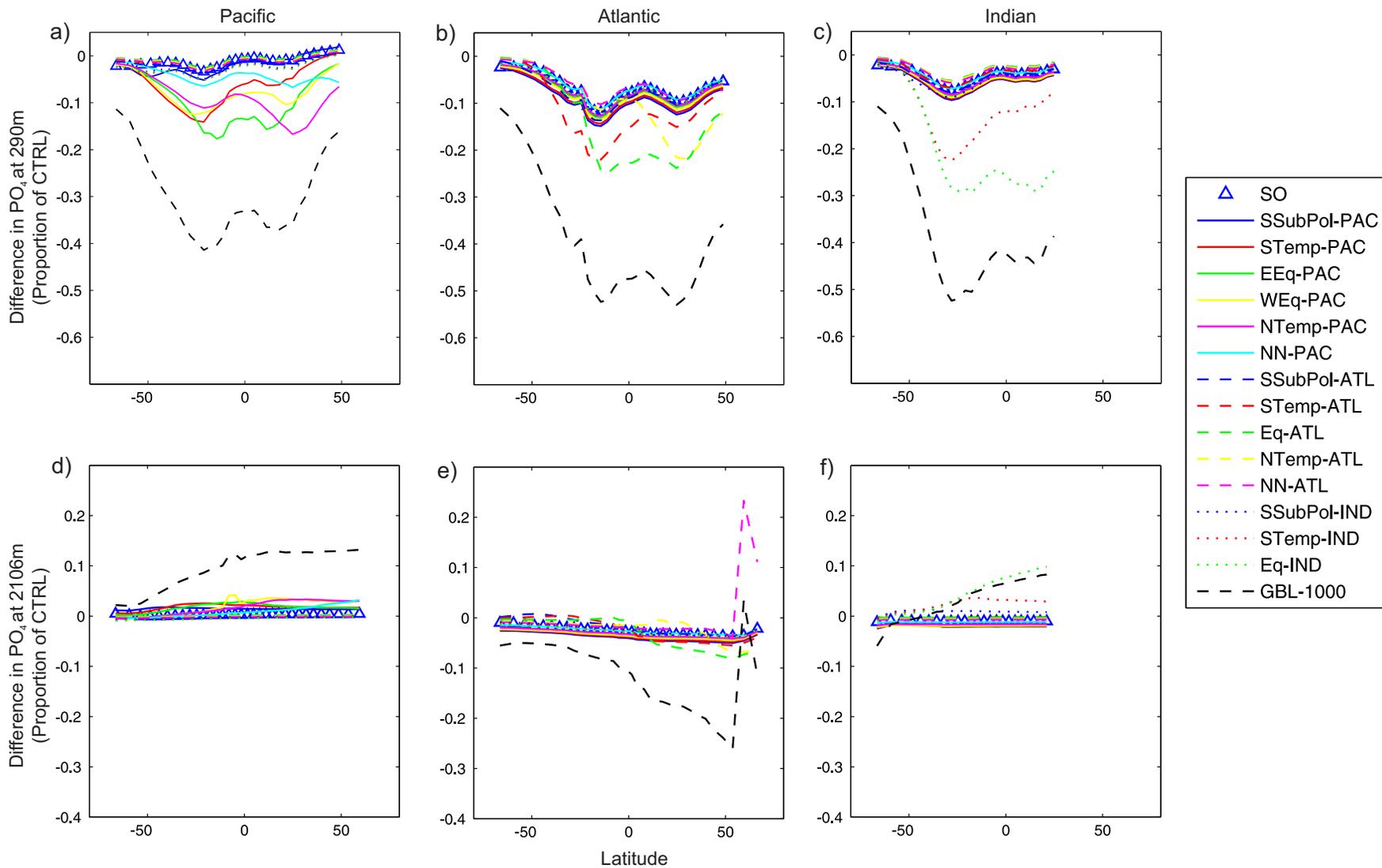
Figure 5.7: See caption for Figure 5.6. Remineralisation depths are now changed to 1000m. The layout and formatting are identical to Figure 5.6.

For the Pacific and Indian Oceans, PO$_4$ concentrations follow a general rule that a deepening of remineralisation depths will increase the PO$_4$ concentration of the deep ocean and decrease the surface [PO$_4$], *e.g., Boyd and Trull* (2007). However, the deep Atlantic displays an inverse trend (Figures 5.6e and 5.7e). The largest differences in deep Atlantic [PO$_4$] occur when remineralisation depths are perturbed in the Eq-ATL and NTemp-ATL regions (green and yellow dashed lines in Figures 5.6e and 5.7e). The concurrent decrease in surface [PO$_4$] in these regions suggests that changes in surface [PO$_4$] are being communicated to the deep Atlantic via deep water formation in the North Atlantic. These surface signals are potentially larger in size than the input of PO$_4$ from remineralisation that the net change in [PO$_4$] reflects changes at the surface. Similar variability in Atlantic [PO$_4$] was observed by *Kwon and Primeau* (2006) and *Kwon et al.* (2009). In the NN-ATL region, where deep water formation is occurring, [PO$_4$] does follow the general rule of thumb, but this is a very localised change (Figures 5.6e and 5.7e). Other factors, such as changes in the magnitude of PO$_4$ particulate export in these regions as well as the circulation patterns could possibly contribute to this feature.

Regional changes in remineralisation depth in the Indian Ocean do approach the same magnitude of change when the global remineralisation depth is perturbed (Figure 5.6c and 5.6f). Elsewhere differences in [PO$_4$] from regional changes are much smaller in magnitude. In particular, changes in Southern Ocean and high latitude [PO$_4$] are relatively small across all depths, basins and regions where remineralisation changes occur. Equally, changes in the remineralisation depth in the Southern Ocean have a minimal impact on other regions (compare triangles to circles in Figures 5.6 and 5.7).

To explore the reasons behind the different sensitivies in each region, the [PO$_4$] fields at the end of each sensitivity run are compared to the CTRL run [PO$_4$] using the RMSD (Equation 5.6) as a measure of how different the resulting [PO$_4$] field is. This is then compared with the total export flux of PO$_4$, the fractional change in the export flux and the standard deviation of the mean vertical profile of [PO$_4$] in each region

to explore why [PO$_4$] is sensitive to remineralisation changes in some regions and not others (Figure 5.8). There are strong positive correlations between RMSD and the total PO$_4$ export flux in each region (r=0.79 and r=0.90; Figure 5.8a and 5.8d). It follows that the perturbations to remineralisation depth in regions with the largest export flux of PO$_4$ result in the largest perturbations to the global [PO$_4$] distribution. A notable outlier to this relationship is the Southern Ocean (red stars in Figure 5.8) that has a much lower impact on global [PO$_4$] despite the large magnitude of the export flux in this region. This suggests there are other factors affecting the sensitivity to changes in remineralisation depth. Although the magnitude of the export flux is important, changes in remineralisation depth can change the magnitude of the export flux. The fractional change in the export flux is an indication of how the magnitude of the export flux responds to changes in the remineralisation depth (Figure 5.8b and 5.8e). Although there is a correlation (r=0.53 and r=0.67), the relationship is less clear. There is a clear grouping of high-latitude regions with small changes in export that, in the model, are limited by light and less responsive to changes in nutrient levels. In regions that are not limited, there is no clear relationship (Figure 5.8e). Finally, the standard deviation of the mean profile of [PO$_4$] in each region from the CTRL run is used as a measure of water column homogeneity and vertical mixing. The strong positive correlations suggest that changes in the remineralisation depth in vertically homogenous regions, such as the Southern Ocean, have a minimal impact on [PO$_4$] distributions (r=0.65 and r=0.76; Figure 5.8c and 5.8f). Vertical mixing therefore also has an important role in setting the sensitivity of [PO$_4$] to regional changes in remineralisation depth.
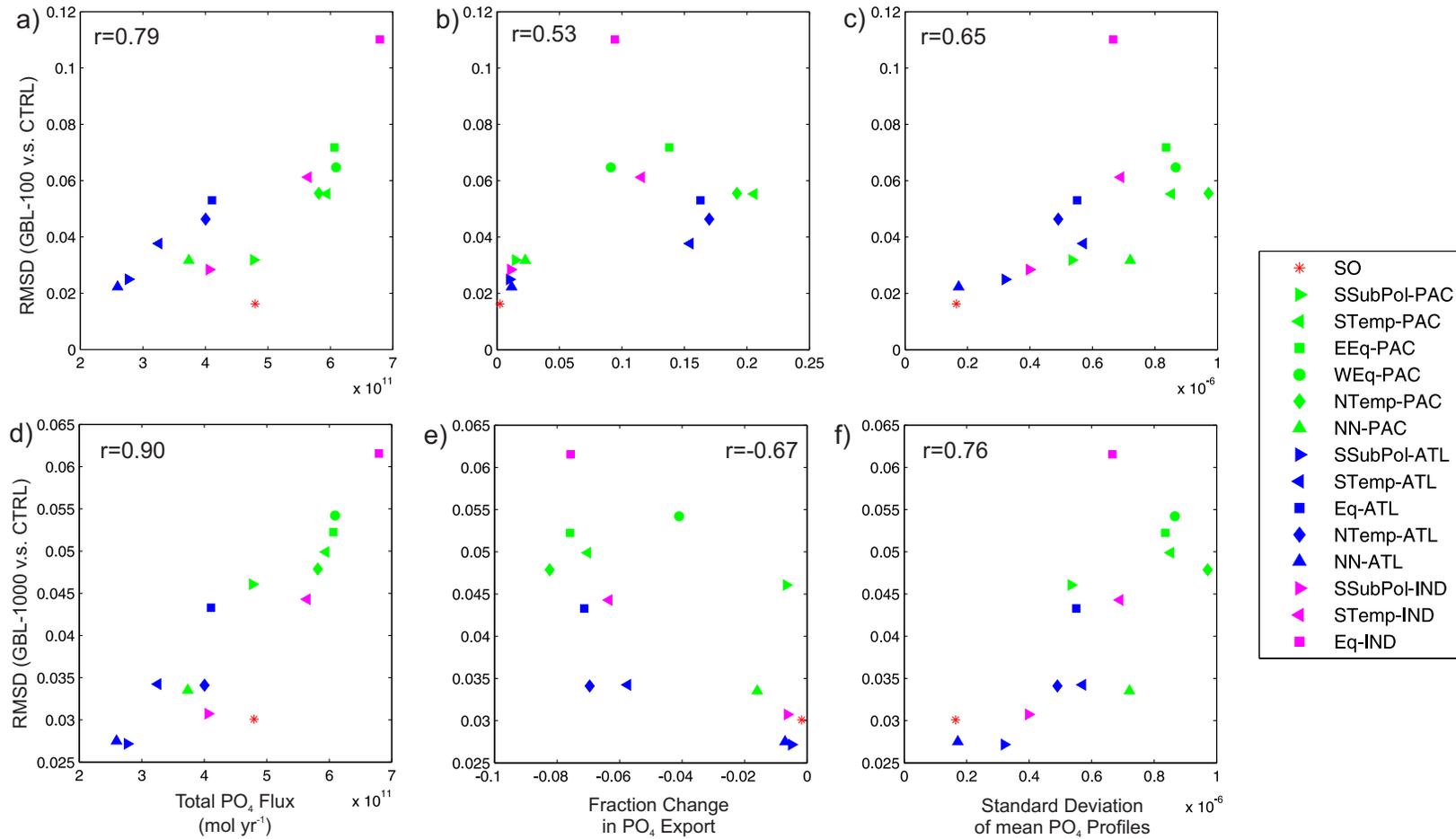
Figure 5.8: Comparison of the difference between the global [PO$_4$] from each sensitivity experiment (SENS-100; SENS-1000) to the control (CTRL) and various metrics for each region. The RMSD is used to calculate a misfit between each experiment and the control [PO$_4$]. Results are shown for both the shallow remineralisation depth (SENS-100, panels a,b and c) and deeper remineralisation (SENS-1000, panels d,e and f). Panels (a) and (c) compare the total export PO$_4$ flux from each region, panels (b) and (d) compare the fractional change in PO$_4$ export, and panels (c) and (e) compare the standard deviation of the mean water column profiles of [PO$_4$] in each region.

5.4   Optimal Solutions of Spatially Variable Remineralisation Depths

After perturbing the remineralisation depth in each region individually, the remineralisation depths in all regions are varied simultaneously using the Latin hypercube approach. A total of 600 individual simulations are run, each with a different set of remineralisation length scales sampled to cover a range of different parameter combinations (see Figure C.4). The global mean remineralisation depths from each of the 600 runs range from 334m to 731m. The Latin hypercube is unlikely to sample all regions towards the minimum or maximum because of the constraints and its requirement to fill the parameter space (see Figure 5.5). The 600 runs are summarised on a Taylor diagram (*Taylor*, 2001) with the global mean remineralisation depth of each run shown on a colour scale (Figure 5.9). The runs span a large part of the range between the GBL-100 and GBL-1000 experiments spanning a range of RMSD values suggesting the LHS ensemble is representative and reliable.

The statistical fit to observations is strongly affected by the global mean remineralisation depth with the LHS experiments moving towards the GBL-100 and GBL-1000 experiments. Runs that are an improvement on the CTRL run, *i.e.,* that have a better RMSD score, are clustered around the CTRL run and the global mean remineralisation depth of 550m. Correlation coefficients for all the runs are relatively similar suggesting changes in the remineralisation depth do not significantly change the overall structure of the [PO$_4$] field suggesting that it may be primarily determined by the ocean circulation (this can also be seen in the similar shapes of the zonal averages in Figures 5.6 and 5.7). The global normalised variability of [PO$_4$] is, however, more sensitive to remineralisation, but is primarily related to the global average remineralisation depth rather than regional changes.
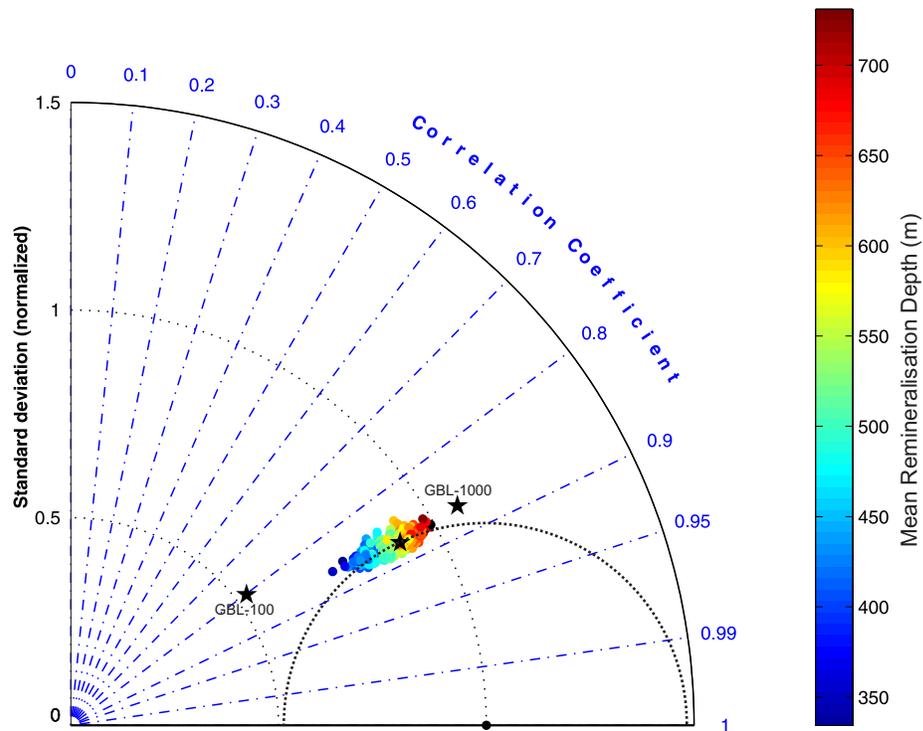
Figure 5.9: Pattern statistics of all 600 runs compared to World Ocean Atlas observed [PO$_4$] (coloured markers) from the LHS experiment plotted as a Taylor diagram. The radial axis indicates the correlation between the model and observation fields, whilst the linear axis indicates the standard deviation, *i.e.,* the global variability of [PO$_4$] normalised to the standard deviation of the observations. The black marker on the lower axis indicates a perfect match. The CTRL run is shown as the black marker in the middle of the data. The stars indicate GBL-100 and GBL-1000 experiments. The colour of the LHS experiments indicates the global mean remineralisation depth for that run.

5.4.1   Spatial Variability in Remineralisation Depths

A set of improved model runs is defined from the LHS experiments as those with a RMSD score lower than the CTRL run (RMSD<0.4932). In total, 155 runs out of the 600 are defined as a better fit than the CTRL, with a minimum RMSD of 0.4206 being achieved. Each parameter has a distribution of 155 values which are displayed using a boxplot (Figure 5.10). The boxplot has the advantage of summarising non-parametric distributions that can be irregularly shaped and that contain outliers (*Krzywinski and Altman*, 2014), a characteristic of the random sampling of Latin hypercube samples. A small box (representing the interquartile range) indicates a well constrained parameter,

i.e., there is a well-defined range of parameter values across the set of best model runs. If the box is large, this may indicate either a parameter that is unconstrained, *i.e.,* the value has no bearing on the fit to data, or that it may be covarying with another parameter(s), *i.e.,* the parameter is not independent. The boxplots also indicate 95% confidence intervals (CIs) for each box (triangle markers). In general, when a set of CIs do not overlap, the medians can be judged to be significantly different (*Krzywinski and Altman*, 2014).

The box plots in Figure 5.10b and 5.10c are arranged by basin and increasing latitude so as to correspond with the predominantly latitudinal variability seen in sediment trap observation (Figure 5.10a). Since the subset of runs is fairly large, a smaller subset of the best 40 runs is also shown as a check that any features are consistent across the runs (Figure 5.10c). Subsets smaller than 40 result in CIs that are larger than the IQR suggesting they are not reliable (*Krzywinski and Altman*, 2014). Both Figure 5.10b and Figure 5.10c show similar patterns suggesting that they are a real feature of the dataset and not a result of choosing subsets of data. The Atlantic displays a significant latitudinal relationship that is consistent with sediment trap observations (deeper in low latitudes and shallower at high latitudes). The Equatorial Atlantic region has a significantly deeper median length scale than the CTRL (709-824m *cf.* 550m). In the Pacific, the northern high latitude regions also have a significantly shallow length scale (182-348m for the North Pacific, 289-364m for Temperate North Pacific). However, the latitudinal trend south of the equator is less distinct. The East, West Equatorial Pacific and Temperate South Pacific regions are not significantly distinguishable. The SubPolar Pacific region is significantly deeper on average than the other Pacific regions. Finally, the Southern Ocean is very similar to the CTRL value with a large box relative to the median in Figure 5.10, suggesting the remineralisation depth in this region cannot be constrained or is not independent. The Indian Ocean regions display similar features to the Southern Ocean. Similar spatial patterns for the optimal solutions are found when using higher resolution regions (Figure C.8).
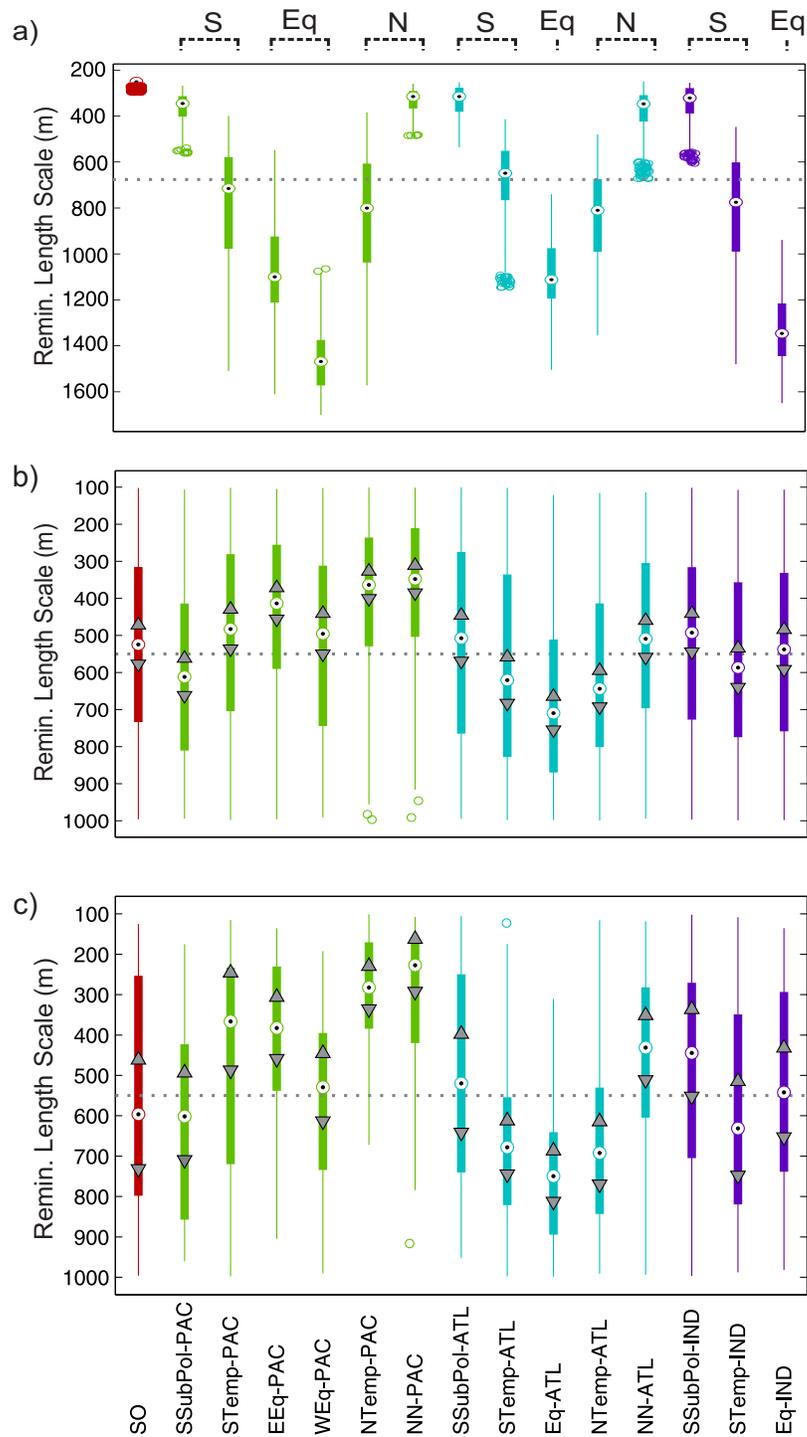
Figure 5.10: Boxplots showing the distribution of remineralisation depths in each region from observations and the best runs from the LHS experiments. (a) boxplots of *e*-folding depths calculated from the Martin Curves in *Henson et al.* (2012). (b) all LHS experiments with a smaller RMSD than CTRL (n=155), and (c) a subset of the 155 runs with the smallest RMSD (n=40). The target indicates the median of each distribution, the box indicates the interquartile range, and the triangles indicate confidence intervals that are used to judge if the medians are significantly different from each other. Boxes are arranged by basin and latitude. The dashed horizontal line rendicates in (b) and (c) indicate the CTRL remineralisation depth of 550m and the mean *e*-folding depth from *Henson et al.* (2012). Circles indicate outliers defined as outside 1.5 x interquartile range.   140

The reasons for the unconstrained regions can be explored further using the distribution of the parameters values for each region indicating that some of the large boxes in Figure 5.10 are potentially the result of multiple optimums (Figure 5.11). For example, the Southern Ocean, Temperate South Pacifc and SubPolar Atlantic regions have distributions that may indicate multiple optimum solutions for the remineralisation depth in those regions. The pattern of remineralisation depths in Figure 5.11 can be related to the sensitivity results in Figure 5.8. For example, global $[PO_4]$ was shown to be relatively insensitive to remineralisation changes in the Southern Ocean, possibly due to the extent of vertical mixing. This corroborates the large box size in Figure 5.10 suggesting that this region is difficult to constrain using $[PO_4]$. In contrast to Figure 5.8, the Indian Ocean is not constrained despite its large impact on global $[PO_4]$, whilst the Northern Pacific is well constrained despite its small impact.

To test whether there are regions that are covarying with each other as suggested by the distributions in Figure 5.11, partial correlations are calculated between the remineralisation depths of the best 40 runs between each region (Figure 5.12). As all regions are varied simultaneously, a standard correlation would not be able to separate the variability between two regions and other regions. Significant negative correlations exist between the Southern Ocean and the Sub Polar and Temperate South Pacific. This suggests that changes in remineralisation depths in the Sub Polar and Temperate South Pacific are compensating for concurrent changes in the Southern Ocean. This corroborates the shape of the distributions in Figure 5.11. Equally, this can be seen with the Temperate North Pacific and Equatorial Pacific regions, as well as the Indian Ocean regions. The only positive correlation between regions occurs between the North Atlantic and the Sub Polar Indian Ocean. The reason for this interaction is less clear, although this may reflect an interaction between the Atlantic deep water properties and the Sub Polar Indian Ocean.
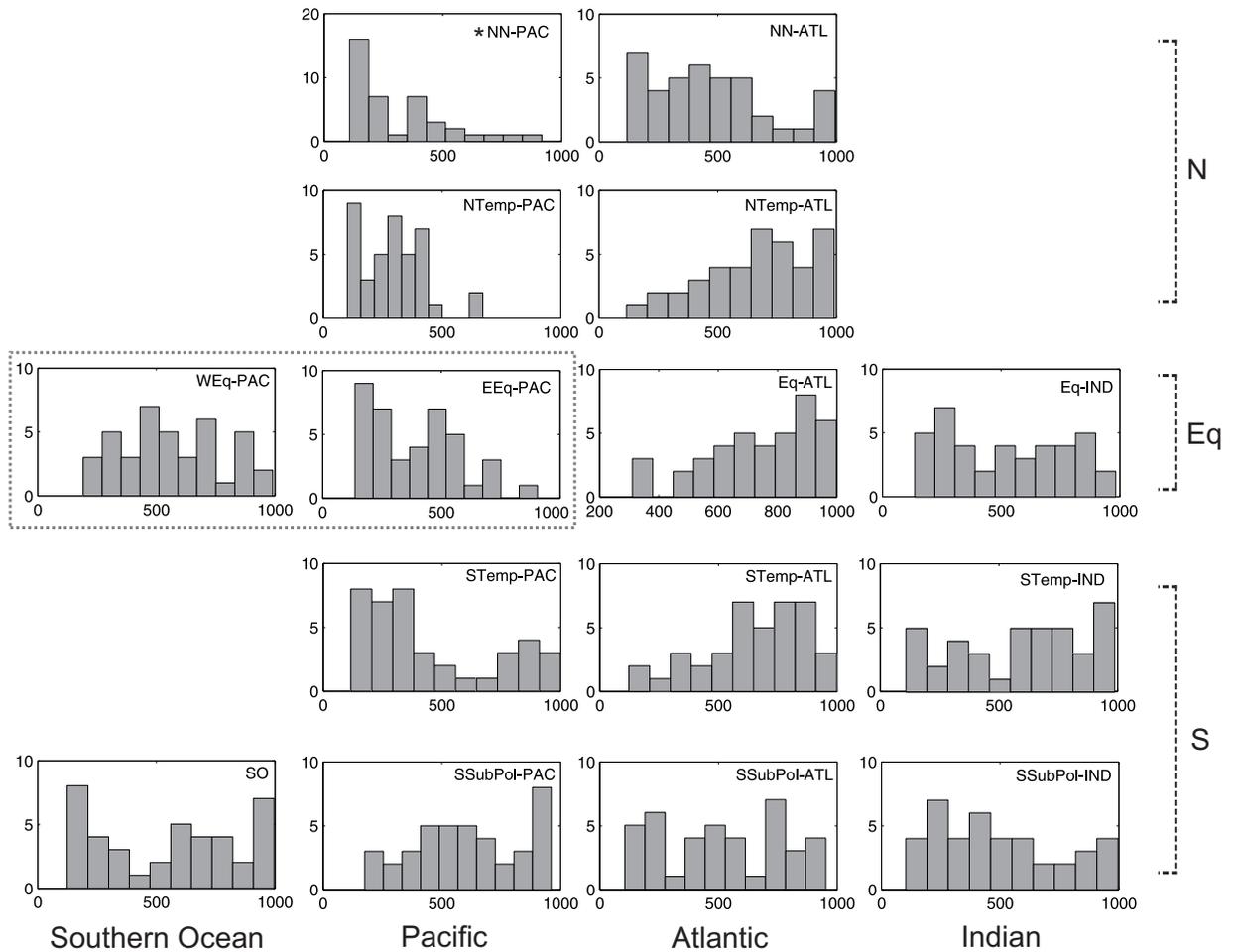
Figure 5.11: Histograms of the parameter distributions from Figure 5.10b. Panels are arranged by basin and latitude. Note the scale difference between the NN-PAC panel and other panels, denoted by *. The dashed box indicates the two equatorial regions in the Pacific.
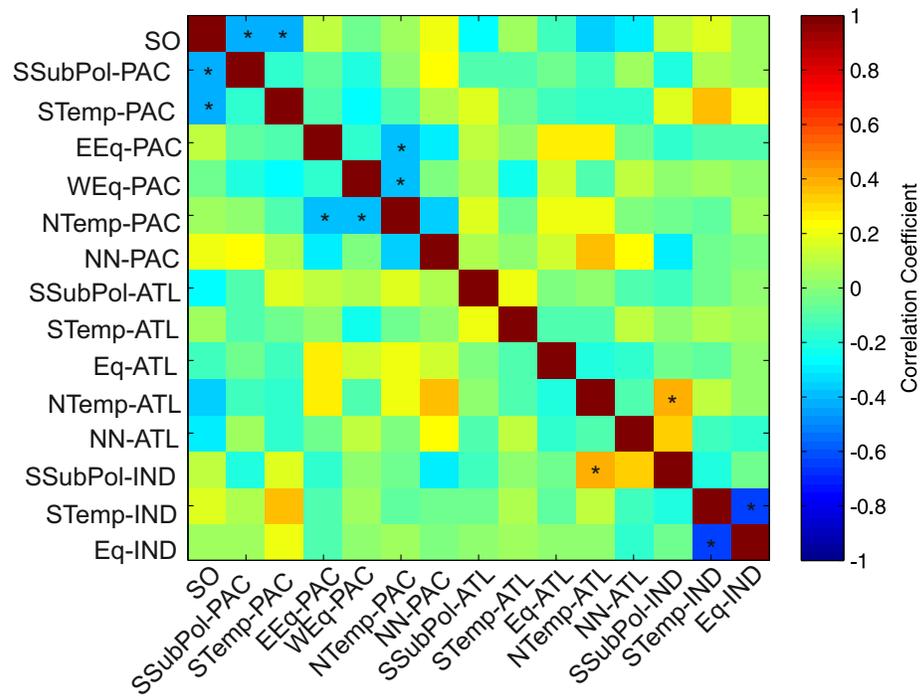
Figure 5.12: Partial correlation coefficients between remineralisation depths in all regions for the 40 best runs of the LHS experiment (see Figure 5.10b). Each partial correlation coefficient calculated from the remineralisation depths between two regions while accounting for the simulataneous variability in all other regions. Stars indicate significance values greater than 95%.

5.5   Discussion

5.5.1   Robustness of the Model and Experiments

A set of experiments have been designed to explore the sensitivity of model $[PO_4]$ distributions to regional changes in remineralisation depths. A combination of spatially variable remineralisation depths can be defined as a best fit to $[PO_4]$ observations, and also assessed for how robust this solution is. The Latin hypercube approach used to achieve this result requires a large number of simulations to explore the parameter space. Two features of this study make this approach feasible: a relatively simple biogeochemical model with a relatively low resolution steady state circulation and the use of pre-defined regions. The extent to which these might introduce biases in the results is explored in the following sections.

5.5.1.1   Model Structure

Previous studies of the sensitivity of model [PO_4] fields to different complexities of biogeochemical models have shown that increased model complexity does not necessarily result in increased model fidelity (*Kriest et al.*, 2010 and *Kriest et al.*, 2012; the biogeochemical model used here is equivalent to the N+DOP model used in these studies). Both studies show that the global [PO_4] field is sensitive to changes in the global remineralisation depth regardless of the biogeochemical model used. However, there are number of features not considered that may have an impact on the results in this chapter. The global export parameters used here ($u_0^{PO_4}$ and K^{PO_4}) in equation 5.3) represents a compromise between allowing high surface [PO_4] to be maintained in 'High Nutrient Low Chlorophyll' regions and depleting [PO_4] elsewhere (*Ridgwell et al.*, 2007a). The limitation of nutrient uptake is potentially an important factor in setting the sensitivity of [PO_4] to changes in the biological pump in certain areas (Figure 5.8). There is no explicit representation of iron limitation either in the model used here or in *Kriest et al.* (2010, 2012) which would likely change uptake rates in these areas, such as the Southern Ocean, relative to others. In other optimisation studies a nutrient-restoring scheme has been used for estimating the magnitude of PO_4 export (*Kwon and Primeau*, 2006; *DeVries et al.*, 2014). As this scheme diagnoses export in every surface grid-box separately, regions may respond differently. There is no direct comparison between these two types of schemes to gauge what effect this may have or how this may affect sensitivity and optimisation approaches.

The model used here includes a refractory fraction of sinking organic matter that penetrates much deeper into the ocean interior in contrast with some previous studies that considered only one labile fraction (*Kwon and Primeau*, 2006; *Kriest et al.*, 2012). The inclusion of a refractory fraction has been shown to improve the fit of model nutrient distributions to observations, as single fraction flux curves can underestimate fluxes at depth (*Andersson et al.*, 2004; *Kriest and Oschlies*, 2013; *DeVries et al.*, 2014). A re-

fractory flux has also been combined with a burial flux of PO_4 in sediments although its effect on dissolved $[PO_4]$ distributions is limited by the relatively long residence time of PO_4 in the ocean compared to the timescales of ocean overturning (*Kriest and Oschlies*, 2013). *DeVries et al.* (2014) compared single fraction flux curves, with the inclusion of a second refractory fraction and the inclusion of burial finding that the inclusion of a refractory flux was the most important to fitting $[PO_4]$ observations. Overall, the structure of the biogeochemical model used here sits at the base of the range of models used in similar studies (see also Chapter 2).

In contrast to the biogeochemical model, the coarse resolution of the circulation model will also impart some significant limitations on the analysis in this chapter. For example, the stability of high latitude water columns may be sensitive to the vertical resolution of models (*Müller et al.*, 2006). *Ridgwell et al.* (2007a) note that this may be a limitation to the 8-level model used here but that excessive high-latitude ventilation would result in overestimates of PO_4 export that are not observed in the model. The vertical resolution may also introduce numerical problems when parameterising the vertical profile of organic matter remineralisation (*Kriest and Oschlies*, 2011). Also, the steady state circulation used here is not seasonal, which again may be an important factor in simulating the biological pump in the high latitudes. An important next step, therefore, would be to apply the methodology of this study using a model with a higher vertical resolution and/or seasonality. The 16 level version of GENIE forced both seasonally and non-seasonally (*Marsh et al.*, 2011), would be an ideal candidate to test the effects of both resolution and seasonality, although the method of diagnosing a transport matrix used here would need to be adapted for the seasonal circulation.

5.5.1.2   Regions

To facilitate the Latin hypercube approach, the model domain was split into 15 regions by latitude and basin. Although these have been defined using latitude bands and

not dynamically using metrics such as upwelling, Figure 5.4b demonstrates that they are, at least to a first order, consistent with general surface features of [PO$_4$] from upwelling/downwelling and PO$_4$ uptake. Previous studies have suggested that variability in remineralisation is to some extent related to large scale biogeochemical differences such as regions dominated by CaCO$_3$ export or opal export (*Klaas and Archer*, 2002; *Henson et al.*, 2012; *Wilson et al.*, 2012). Therefore, the 15 regions should be able to capture some of this variability. Increasing the number of regions could be an option as the computational cost using this model is still relatively small. However, the results of using a larger number of parameters will become harder to analyse because of the difficulty in dealing with a higher number of dimensions. For example, the features observed in the Southern Ocean and Sub Polar regions in Figures 5.11 and 5.12 suggest that these regions do not have an independent effect on [PO$_4$] via changes in the remineralisation depth. Increasing the number and resolution of regions is likely to exacerbate this problem.

5.5.2   Contrasting the Sensitivity of PO$_4$ to Global and Regional Remineralisation Changes

In previous model optimisation studies, the global [PO$_4$] distribution was found to be more sensitive to changes in the global remineralisation depth than changes in other parameters such as uptake rates and DOM remineralisation (*Kwon and Primeau*, 2006; *Kriest et al.*, 2010, 2012). Similarly in this study, there were large differences between the control run (CTRL) and the two experiments where the remineralisation depth was changed globally (GBL-100 and GBL-1000) (Figure 5.10). In contrast to this, the distribution of global [PO$_4$] and the fit to observations were relatively insensitive to regional changes in remineralisation depths (Figure 5.10).

The sensitivity analysis (Figures 5.6 and 5.7) suggests that changes in [PO$_4$] from changes in the remineralisation depth are relatively local and therefore do not change

the overall fit to observed [PO$_4$]. Although there are differences between regions due to the magnitude of the organic matter export flux and vertical mixing (Figure 5.8), none of the regions significantly contributed to the change in global [PO$_4$] distribution when remineralisation depths were changed globally (Figures 5.6 and 5.7). This suggests that the changes in regional remineralisation are cumulative such that a net export of PO$_4$ from the Atlantic to the Pacific observed when deepening remineralisation (*Kwon and Primeau*, 2006; *Kriest et al.*, 2012) is primarily due to large-scale changes between basins and not between specific regions. The global [PO$_4$] distribution may therefore be relatively constrained by large-scale circulation patterns, rather than by large-scale features in biogeochemistry. This is further supported by the fact that the majority of changes in the model [PO$_4$] field are related to variability (changes in the standard deviation) rather than the spatial structure (changes in the correlation coefficient) (Figure 5.9). If true, this has implications for biogeochemical modelling. For example, this result may suggest that using a global mean remineralisation depth, *e.g.,* the Martin Curve, may be preferable over a spatially explicit parameterisation as it can be constrained. This result may however reflect the similarity in the residence time of PO$_4$, *e.g.,* 1500 years (*Paytan and McLaughlin*, 2007), compared to overturning circulation timescales, *e.g.,* a mean transit time from the surface to deep North Pacific of 1360±350 years (*Khatiwala et al.*, 2012). This is not necessarily true for all related tracers such as oxygen (*Kriest and Oschlies*, 2013).

Another factor that may help to explain why the global [PO$_4$] distribution and fit to observations are insensitive to regional remineralisation changes may be the interactions between preformed and regenerated [PO$_4$]. Preformed nutrients reflect nutrients in the ocean interior that have been advected from the surface, unlike regenerated nutrients that have entered the ocean interior via the remineralisation of sinking particles (*Ito and Follows*, 2005). Similar total [PO$_4$] distributions have been found to arise from different fields of preformed and regenerated [PO$_4$] suggesting that, to some extent, they may act to compensate each other (*Duteil et al.*, 2012). In experiments performed

here, the Atlantic was found to be more sensitive to changes in preformed [PO_4] than other regions (Figures 5.6 and 5.7). This could form an additional constraint, although methods of estimating regenerated [PO_4] from observations can be associated with significant spatial errors (*Ito et al.*, 2004; *Duteil et al.*, 2013). This is explored in more detail in section 5.5.4.

5.5.3   Why is there Spatial Variability in the Optimised Remineralisation Depths?

Given that the global distribution of [PO_4] is relatively insensitive to regional changes in remineralisation, one might expect it to be difficult to find an optimal solution(s). However, Figure 5.10 shows that the remineralisation depth in some regions, such as the North Pacific, are well constrained, and that significant latitudinal trends exist across the Pacific and Atlantic. There are some similarities with current hypotheses such ballasting. The Sub Polar Pacific is a region that has previously been shown to have a higher PIC:POC export ratio than in other basins (*Sarmiento et al.*, 2002; *Jin et al.*, 2006). The deeper remineralisation here could reflect a ballast mechanism. However, other regions where carbonate export is relatively high such as the North Pacific do not support this. There are similarities between our results and the sediment trap analysis of *Henson et al.* (2012), particularly in the Atlantic Ocean. This may corroborate their findings suggesting that there are key biogeochemical drivers of the remineralisation of sinking organic matter. For example, our results could tentatively be interpreted as evidence for surface ecosystems playing a key role, due to packaging processes or the lability of organic matter (*Lam et al.*, 2011; *Henson et al.*, 2012). Ultimately, these processes have similar spatial distributions and may not be easily distinguishable from each other (*Lima et al.*, 2014).

An alternative explanation, given the range of previous estimates (Figure 5.2) and the relative insensitivity of [PO_4] to regional remineralisation changes, is that the optimised

remineralisation depths are compensating for deficiencies in the ocean model. *Kriest et al.* (2010) found that the North Pacific was responsible for a large part of the global misfit between model [PO_4] and observations because this region is at the end of the thermohaline circulation and accumulates [PO_4]. The significant shallow North Pacific remineralisation depths found could be responding changes in the deep North Pacific. Nutrient trapping is a key issue for ocean models in this region (*Najjar et al.*, 2007; *Dietze and Loeptien*, 2013). The model may readily accumulate PO_4 in the deep North Pacific, not returning it to the surface. The shallower North Pacific remineralisation may then be compensating for this feature. Interpreting the optimisation results requires further knowledge of these effects before the spatial patterns can be attributed to actual biogeochemical processes.

5.5.4   Implications for Modelling

The experiments in this chapter have shown that overall, the distribution of global [PO_4] in ocean biogeochemical models is relatively insensitive to regional changes in remineralisation. In some regions, differences in remineralisation depth can be constrained, but it is difficult to be certain whether this actually reflects underlying mechanisms or deficiencies in the ocean model. This suggests that global [PO_4] observations may not be a reliable constraint for the spatial variability of remineralisation in ocean biogeochemical models and may explain the range of solutions in Figure 5.2.

It is important to note that although that [PO_4] is relatively insensitive to differences in remineralisation depths, it does not necessarily mean that the uptake of atmospheric CO_2 via the biological pump is also insensitive. To highlight this, the range of global P* values (equation 5.7) are calculated for each set of sensitivity experiments (SENS-100 and SENS-1000) to suggest how sensitive the uptake of atmospheric CO_2 is to changes in remineralisation depth in each region (Figure 5.13). The preformed PO_4 for each sensitivity run is calculated by initialising the surface ocean with the steady-state

surface PO_4 concentrations at the end of each run. The model is then integrated forward in time for 10000 years with the surface $[PO_4]$ being restored to initial surface values. Regenerated $[PO_4]$ (PO_4^{REG}) is calculated as the difference between the total $[PO_4]$ and preformed $[PO_4]$. Alternative methods using apparent oxygen utilisation could be used to estimate regenerated and preformed $[PO_4]$ but are not available in this phosphorous only model.

$$P^* = \frac{\overline{PO_4^{REG}}}{\overline{PO_4}} \tag{5.7}$$

P^* reflects the ratio between preformed $[PO_4]$ and total $[PO_4]$ in the global ocean (*Ito and Follows*, 2005). An increase in P^* is associated with an increase in $[PO_4]$ that is utilised by biology at the surface, and so reflects a more efficient biological pump. Model experiments have shown a near linear relationship between P^* and atmospheric CO_2 (*Ito and Follows*, 2005; *Marinov et al.*, 2008). To estimate the effect of remineralisation depth changes on atmospheric CO_2, a linear fit between modelled P^* and atmospheric CO_2 values found by *Ito and Follows* (2005) is used to predict the resulting CO_2:

$$CO_2 = -416.6734 \cdot P^* + 414.8058 \tag{5.8}$$

Figure 5.13 shows the range of atmospheric CO_2 values predicted from the P^* values when the remineralisation depth is changed to 100m and 1000m. Regions where remineralisation depths were harder to constrain, such as the Southern Ocean, the subpolar South Pacific and the Indian Ocean, are regions where the global uptake of CO_2 is most sensitive. Equally, regions where the remineralisation depth could be constrained have a smaller effect of atmospheric CO_2. Therefore, parameterisations of remineralisation that include explicit spatial variability could result in large differences in carbon cycle feedbacks despite producing similar $[PO_4]$ fields.
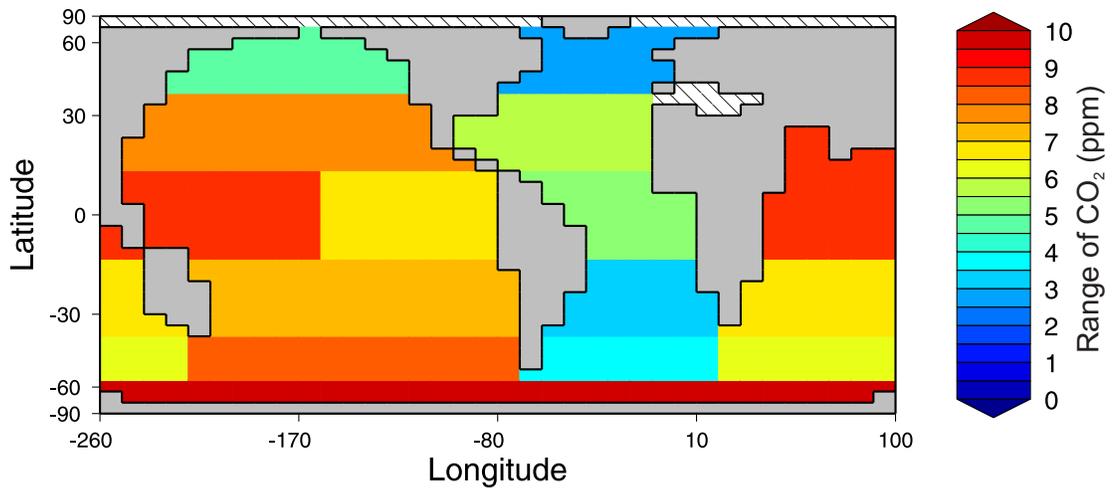
Figure 5.13: Range of CO_2 values in each region when remineralisation depths are changed in each region estimated from P*. P* is calculated from the SENS-1000 experiments and the SENS-100 experiments using equation 5.7 and converted to atmospheric CO_2 via equation 5.8. A larger value indicates that CO_2 uptake via the biological pump is more sensitive in that region. Hatched regions were not tested, see Figure 5.4a.

5.6   Conclusions

In this chapter, the sensitivity and optimisation of regional differences in the remineralisation depth of organic matter have been explored using a new steady-state phosphorus-only version of the Earth system model GENIE and a Latin hypercube approach. Although the sensitivity of model $[PO_4]$ fields to regional changes in the remineralisation depth were variously sensitive to the magnitude of the PO_4 export and vertical mixing, this sensitivity was relatively small when compared with global changes in remineralisation. Regional changes in $[PO_4]$ occurring from remineralisation were predominantly local suggesting the sensitivity to global changes in remineralisation reflect large-scale patterns in ocean circulation rather than in biogeochemistry. An improved fit between model $[PO_4]$ and observations can be achieved, with the resulting patterns showing some similarities to those found in sediment trap studies. However, further work is needed to reliably interpret these as biogeochemical features and not caveats from overfitting the model to data. These findings suggest that $[PO_4]$ is not a good constraint for regional variability in remineralisation depths. The uptake of

atmospheric CO_2 is however sensitive to these changes.

CHAPTER 6

Conclusions and Further Work

The aim of this thesis was to evaluate and advance the current mechanistic understanding of the biological pump in the ocean interior. The particular focus was on exploring the global scale function of the remineralisation of sinking particles in the ocean interior to be able to relate mechanisms back to their effect on atmospheric CO_2 concentrations.

Chapter 1 showed that atmospheric CO_2 concentrations are potentially sensitive to relatively small changes in the depth of remineralisation sinking particles of organic matter. This occurs via the redistribution of DIC and alkalinity as well as changes in the alkalinity inventory driven by changes in the ratio between POC and $CaCO_3$ fluxes at the sediment surface. Global observations of sinking particles, primarily from sediment traps, have revealed significant spatial variability in the depth of POC remineralisation in the modern ocean leading to the development of a number of mechanistic hypotheses. The ballast hypothesis predicts that the sinking velocity of POC is increased by the presence of certain minerals such as $CaCO_3$ which has significant implications for understanding how environmental changes in the past and future may impact particles fluxes and atmospheric CO_2. Alternative interpretations include a more prominent role for ecosystem processes such as the quality of organic matter being exported. The uncertainty in the range of mechanisms is accompanied by recent research questioning the nature of the spatial variability in POC fluxes. This presents a large uncertainty in

how the biological pump may have contributed to past changes in atmospheric CO_2 as well as its response to future climate change.

In response to this, chapters 3 presents data analysis aimed at exploring uncertainties in the ballast hypothesis,whilst chapters 4 and 5 present new approaches to quantifying the biological pump function using an Earth system model. In the rest of this chapter, the main findings and implications of each chapter are reviewed before discussing directions for future work.

6.1   Summary of Results

6.1.1   Spatial Variability in Ballasting Statistics

A role for the ballasting of POC by minerals was inferred from significant global correlations (carrying coefficients) between POC and $CaCO_3$ in deep sediment trap measurements (*Klaas and Archer*, 2002). Regional differences in opal fluxes have been cited as reason why a global analysis may not be appropriate (*De La Rocha and Passow*, 2007). Exploring these uncertainties further requires data to be grouped into representative regions, *e.g. Le Moigne et al.* (2012), but problems arise from groupings that either result in small sample numbers or that are too large that they average any underlying variability. Geographically Weighted Regression (GWR) presents a practical solution to this problem by defining regions in a statistically rigorous and consistent way that allows carrying coefficients to vary in space.

GWR was applied to an updated global sediment trap database. Examination of carrying coefficients at a basin scale, the spatial pattern of residuals from the global regression, and statistics from the GWR analysis suggested there was significant spatial variability in carrying coefficients. The carrying coefficient for $CaCO_3$ was the most variable and shown to be robust. The lack of a consistent strong global relationship between POC and $CaCO_3$ suggests that there is more to the controls on sinking POC

than just the dominant mineral present. Spatial patterns in the $CaCO_3$ coefficients correspond with broad patterns in surface biogeochemistry supporting a role for ecosystem processes in setting vertical POC fluxes. The GWR model predicts that much less POC is associated with $CaCO_3$ suggesting that a ballasting parameterisation in ocean biogeochemical models may overestimate fluxes of POC to depth as well as potentially reflecting inappropriate responses to global change. Further interpretation of the carrying coefficients from the GWR analysis was limited by the relatively low sampling density of sediment traps.

6.1.2   Estimating POC Flux Profiles from Observed Tracers

In chapter 3, a major limitation to linking the spatial variability in carrying coefficients to a mechanistic basis was the sparse sampling of data in space. In response to this problem, chapter 4 explored whether vertical profiles of POC could be derived from high-resolution fields of tracer measurements using modelled circulation rates. This process involves first estimating rates of remineralisation in the water column which represent the first derivative of the flux profile such as the Martin Curve. Flux curves can be inferred by integrating remineralisation rates vertically. When applied to Apparent Oxygen Utilisation (AOU), representing the product of remineralisation integrated over time with ocean circulation, estimates of remineralisation rates are integrated over large areas of the ocean. An alternative approach was explored using the steady-state circulation from a model, diagnosed as a Transport Matrix (TM), to estimate remineralisation rates from observed tracers such as phosphate (PO_4). A synthetic dataset in the Earth System Model GENIE was created to test the uncertainty arising from various sources of error. Due to this approach not being applied on a global scale before, the analysis also explored sources of error when inferring flux curves from remineralisation rates.

Inverting a model [PO_4] field with a TM from the corresponding model circulation re-

produces remineralisation rates. However, the approach is very sensitive to both sources of error associated with the [PO$_4$] observations and from the use of model transport rates. Uncertainty in tracer measurements have a large effect on estimated remineralisation rates due to the way that the TM is constructed. Differences between the modelled and observed transport rates also contribute uncertainty as transport rates are 3-4 orders of magnitude larger than remineralisation rates. Errors from the model transport rates are likely to be systematic suggesting a potential way of constraining the size of these errors although this is still subject to the uncertainties form the observations. Our results also suggest that the presence of DOM remineralisation when inferring particle flux curves from profiles of remineralisation rates can produce spatial patterns similar to observed patterns in particle fluxes (*Henson et al.*, 2012). These results suggest that even if remineralisation rates could be estimated accurately, there is still significant uncertainty when using them to information about particle fluxes.

6.1.3 Optimisation and Sensitivity of Regional Remineralisation Depths

Nutrient distributions and atmospheric CO_2 concentrations are sensitive to changes in the global average depth of POC remineralisation in models. Model optimisation studies have used this sensitivity to find global mean remineralisation depths that result in the statistically best fit to tracer observations such as PO$_4$. This provides an alternative approach to the one described in chapter 4 to exploring spatial variability in POC fluxes. Chapter 5 explored whether there is an optimal set of regionally variable remineralisation depths in the Earth system model that best fits the observed annual average [PO$_4$] field and how robust these solutions are.

A computationally fast phosphorous-only version of the Earth system model GENIE using the TM from chapter 4 was developed to represent steady-state circulation. The ocean was divided into 15 biogeochemical biomes within which the remineralisation depth was an independent parameter. Latin hypercube sampling was used to produce

an ensemble of runs that efficiently sample across the range of potential combinations of remineralisation depths, producing probability distributions for each region. Despite sensitivity to the global remineralisation depth, the distribution of [PO$_4$] is shown to be relatively insensitive to regional changes in remineralisation. An optimal combination of remineralisation depths in the Atlantic is found that predicts deeper remineralisation in the low latitudes and shallower at high latitudes, matching sediment trap observations. Shallow remineralisation is also predicted in the North Pacific. However, remineralisation depths in the Southern Ocean, South and Equatorial Pacific, and Indian Ocean cannot be successfully constrained. The comparatively large sensitivity to global remineralisation depths and insensitivity to regional changes question whether the optimal solution found is significant. Our results suggest that nutrient observations, such as [PO$_4$], may not be able to distinguish reliably between different mechanistic models of the biological pump. A comparison of how sensitive preformed PO$_4$ is in each region also suggests that the regions that could not be constrained are also the ones where changes in the remineralisation depth has the largest effect on the air-sea gas exchange of CO$_2$.

6.2   Future Work

Future work to follow on from the findings presented in this thesis can be broadly grouped into two directions. Firstly, there is still a need to better understand the controls on carbon fluxes in the ocean interior and secondly there is a need to understand the implications of these mechanisms through both better representation in biogeochemical models and assessing uncertainties.

6.2.1   Towards a Mechanistic Understanding

A mechanistic understanding of the key controls on carbon fluxes in the ocean interior is clearly important given the range of current hypothesised mechanisms and the uncer-

tainty in the interpretation of existing observations (Chapter 3). Although deep ocean sediment traps are an important tool, they can only provide a snapshot of the end results of key processes occurring higher up in the water column (*Ragueneau et al.*, 2006). Increasing the number and type of observations in the mesopelagic is a clear starting point to improving our understanding. New techniques to provide direct measurements, such as neutrally-buoyant sediment traps (*Buesseler et al.*, 2007), filtration units to measure particle concentration (*Lam et al.*, In Press) and indirect measurements, such as radioisotopes (*Le Moigne et al.*, 2013), will provide important observations in this key region of the water column. Observations of organic matter in sediments, the ultimate end-point for particle fluxes, could also provide additional constraints on mechanisms, *e.g. Seiter et al.* (2005) and *Arndt et al.* (2013). The output from the GEOTRACES programme is an important development in this area (*Anderson and Henderson*, 2005).

The approaches applied in this thesis have focussed on spatial variability on an annual average timescale. A particular difficulty in this approach is that potential spatial variability from ballasting and ecosystem processes may be very similar. A study aimed at inferring processes from annual particle fluxes simulated in a model using a range of processes found that it was very hard to distinguish between processes because they are inextricably linked (*Lima et al.*, 2014). As well as a focus on the mesopelagic, additional analysis of how fluxes change through time, *e.g.* on an intra-annual timescale, may facilitate the separation of processes. For example, a dominant effect from ecosystem processing may be evident with seasonal changes in POC fluxes at depth. Datasets and analysis by *Lutz et al.* (2007) and *Lam et al.* (2011) provide initial examples of this.

An additional approach to circumventing the problem highlighted by *Lima et al.* (2014) is use of the palaeo-record. Past changes in climate and the ocean can provide us with experiments which can be used to test existing hypotheses against. The data-model comparison described by (*John et al.*, 2014), that work during this thesis contributed to, presents a first step towards this. Global mean temperature during the Eocene epoch (55.5 33.7 Mya) was much warmer than the modern period with largely ice-free

poles with implications for ocean biogeochemistry such as increased biological rates (*Matsumoto*, 2007). A new method developed by *John et al.* (2013) reconstructs vertical gradients of $\delta^{13}C$ from foraminiferal calcite exploiting the fact that individual species calcify at different depths. The resulting water column $\delta^{13}C$ profiles through the Eocene display a much steeper and larger gradient than observed in the modern. Remineralisation of organic matter is thought to be a major control on vertical $\delta^{13}C$ gradients (*Holden et al.*, 2013; *Schmittner et al.*, 2013) suggesting temperature dependent remineralisation may be important. A temperature dependent remineralisation rate was incorporated into a late Paleocene/early Eocene configuration of.the Earth System Model GENIE using an Arrhenius-type function. Modelled $\delta^{13}C$ reproduced key aspects of the reconstructed gradients supporting an interpretation of temperature dependent remineralisation rates (*John et al.*, 2014). The palaeo-record may therefore provide additional constraints on our mechanistic understanding. Past events associated with ocean acidification could prove useful when looking for evidence for the ballast hypothesis (*Hönisch et al.*, 2012).

6.2.2   Towards a Dynamic Understanding

Moving towards a better mechanistic understanding of carbon fluxes in the ocean interior also requires a better representation of these processes in ocean biogeochemical models in order to fully understand the implications of these processes. The findings from chapter 5 suggest that nutrient tracers are more sensitive to global changes in remineralisation depths rather than regional changes. Aside from suggesting that one should be cautious in interpreting good fits as evidence for a particular parameterisation, this also further supports using the palaeo-record as it may be reasonable to expect that there were large-scale changes in remineralisation depths. Given the range of hypothesised mechanisms and current uncertainty over dominant mechanisms, an important avenue for future work is to systematically explore the implications for different mechanistic interpretations on past and future changes in climate.

The representation of remineralisation in the ocean interior has traditionally been modelled using Martin Curves or parameterisations of particular mechanisms such as ballasting or temperature. Moving towards a more mechanistic and explicit modelling of these processes would be an important development. Surface ecosystem dynamics are already incorporated into models such as nutrient-phytoplankton-zooplankton-detritus (NPZD) models (*e.g., Yool et al.*, 2013) and the DARWIN model where plankton diversity and biogeography is an emergent property of the model (*Follows et al.*, 2007; *Dutkiewicz et al.*, 2009). These could form the basis for developing models of ocean interior fluxes that are dependent on some aspect of the surface ecosystem, whether related to a ballasting effect from different phytoplankton groups or an organic matter quality effect. Mechanistic models of organic matter degradation for marine sediments should also be used as a basis for development (*Arndt et al.*, 2013). As such, the flux of carbon in the ocean interior is the link between the surface ecosystem where carbon is fixed and the sediments where carbon is buried and recycled. Developing better models of the processes in the ocean interior in line with these other areas could ultimately facilitate the understanding of how the biological pump fully interacts with the carbon cycle and its role in changing CO_2 and climate.

References

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, *19*(6), 716 – 723, doi:10.1109/TAC.1974.1100705.

Alonso-Gonzalez, I. J., J. Aristegui, C. Lee, and A. Calafat (2010), Regional and temporal variability of sinking organic matter in the subtropical northeast Atlantic Ocean: a biomarker diagnosis, *Biogeosciences*, *7*(7), 2101–2115.

Anderson, L. A., and J. L. Sarmiento (1994), Redfield ratios of remineralization determined by nutrient data analysis, *Global Biogeochemical Cycles*, *8*(1), 65–80, doi:10.1029/93GB03318.

Anderson, R., and G. Henderson (2005), Program update: GEOTRACES - a global study of the marine biogeochemical cycles of trace elements and their isotopes, *Oceanography*, *18*(3), 76 – 79, doi:http://dx.doi.org/10.5670/oceanog.2005.31.

Andersson, J. H., J. W. M. Wijsman, P. M. J. Herman, J. J. Middelburg, K. Soetaert, and C. Heip (2004), Respiration patterns in the deep ocean, *Geophysical Research Letters*, *31*(3), L03,304, doi:10.1029/2003GL018756.

Annan, J., J. Hargreaves, N. Edwards, and R. Marsh (2005), Parameter estimation in an intermediate complexity earth system model using an ensemble kalman filter, *Ocean Modelling*, *8*(1–2), 135 – 154, doi:http://dx.doi.org/10.1016/j.ocemod.2003.12.004.

Archer, D. (1991), Modeling the calcite lysocline, *Journal of Geophysical Research: Oceans*, *96*(C9), 17,037–17,050, doi:10.1029/91JC01812.

Archer, D., and E. Maier-Reimer (1994), Effect of deep-sea sedimentary calcite preservation on atmospheric CO_2 concentration, *Nature*, *367*(6460), 260–263, doi:10.1038/367260a0.

Archer, D., et al. (2009), Atmospheric lifetime of fossil fuel carbon dioxide, *Annual Review of Earth and Planetary Sciences*, *37*(1), 117–134, doi:10.1146/annurev.earth.031208.100206.

Armstrong, R., M. Peterson, C. Lee, and S. Wakeham (2009), Settling velocity spectra and the ballast ratio hypothesis, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*(18), 1470 – 1478, doi:http://dx.doi.org/10.1016/j.dsr2.2008.11.032, medFlux:Investigations of Particle Flux in the Twilight Zone.

Armstrong, R. A., C. Lee, J. I. Hedges, S. Honjo, and S. G. Wakeham (2002), A new, mechanistic model for organic carbon fluxes in the ocean based on the quantitative association of POC with ballast minerals, *Deep-Sea Research Part. II*, *49*(1-3), 219–236.

Arndt, S., B. Jorgensen, D. LaRowe, J. Middelburg, R. Pancost, and P. Regnier (2013), Quantification of organic matter degradation in marine sediments: A synthesis and review, *Earth-Science Reviews*, *123*, doi:http://dx.doi.org/10.1016/j.earscirev.2013.02.008.

Barker, S., J. A. Higgins, and H. Elderfield (2003), The future of the carbon cycle: review, calcification response, ballast and feedback on atmospheric CO_2, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *361*(1810), 1977–1998, doi:10.1098/rsta.2003.1238.

Behrenfeld, M. J., and P. G. Falkowski (1997), Photosynthetic rates derived from satellite-based chlorophyll concentration, *Limnology and Oceanography*, *42*(1), 1–20, doi:10.4319/lo.1997.42.1.0001.

Bendtsen, J., K. Hilligse, J. Hansen, and K. Richardson (2015), Analysis of remineralisation, lability, temperature sensitivity and structural composition of organic matter from the upper ocean, *Progress in Oceanography*, *130*(0), 125 – 145, doi:http://dx.doi.org/10.1016/j.pocean.2014.10.009.

Berelson, W. (2001), The flux of particulate organic carbon into the ocean interior: a comparison of four U.S. JGOFS regoinal studies, *Oceanography*, *14*(4), 59 – 67.

Boyd, P. W., and T. W. Trull (2007), Understanding the export of biogenic particles in oceanic waters: Is there consensus?, *Progress in Oceanography*, *72*(4), 276–312, doi:10.1016/j.pocean.2006.10.007.

Broecker, W. S. (1982), Glacial to interglacial changes in ocean chemistry, *Progress in Oceanography*, *11*, 151 – 197.

Broecker, W. S., et al. (1998), How much deep water is formed in the Southern Ocean?, *Journal of Geophysical Research: Oceans*, *103*(C8), 15,833–15,843, doi:10.1029/98JC00248.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones (2006), Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850, *Journal of Geophysical Research: Atmospheres*, *111*(D12), doi:10.1029/2005JD006548.

Brunsdon, C., A. S. Fotheringham, and M. E. Charlton (1998), Geographically weighted regression - modelling spatial non-stationarity, *The Statistician*, *47*(3), 276 – 312, doi:10.1111/1467-9884.00145.

Brzezinski, M. A., and D. M. Nelson (1995), The annual silica cycle in the sargasso sea near bermuda, *Deep Sea Research Part I: Oceanographic Research Papers*, *42*(7), 1215 – 1237, doi:http://dx.doi.org/10.1016/0967-0637(95)93592-3.

Buesseler, K., and P. Boyd (2009), Shedding light on processes that control particle export and flux attenuation in the twilight zone of the open ocean, *Limnology and Oceanography*, *54*(4), 1210–1232, doi:10.4319/lo.2009.54.4.1210.

Buesseler, K. O., et al. (2007), An assessment of the use of sediment traps for estimating upper ocean particle fluxes, *Journal of Marine Research*, *65*(3), 345–416.

Burd, A. B., and G. A. Jackson (2009), Particle aggregation, *Annual Review of Marine Science*, *1*(1), 65–90, doi:10.1146/annurev.marine.010908.163904.

Cameron, D. R., T. M. Lenton, A. J. Ridgwell, J. G. Shepherd, R. Marsh, and A. Yool (2005), A factorial analysis of the marine carbon cycle and ocean circulation controls on atmospheric CO_2, *Global Biogeochemical Cycles*, *19*(4), GB4027, doi:10.1029/2005GB002489.

Cao, L., G. Bala, and K. Caldeira (2012), Climate response to changes in atmospheric carbon dioxide and solar irradiance on the time scale of days to weeks, *Environmental Research Letters*, *7*(3), 034,015.

Chikamoto, M. O., A. Abe-Ouchi, A. Oka, and S. L. Smith (2012), Temperature-induced marine export production during glacial period, *Geophysical Research Letters*, *39*(21), L21,601, doi:10.1029/2012GL053828.

Claussen, M., et al. (2002), Earth system models of intermediate complexity: closing the gap in the spectrum of climate system models, *Climate Dynamics*, *18*(7), 579–586, doi:10.1007/s00382-001-0200-1.

Conte, M., R. Ralph, and E. Ross (2001), Seasonal and interannual variability in deep ocean particle fluxes at the Ocean Flux Program (OFP)/Bermuda Atlantic Time Series (BATS) site in the western Sargasso Sea near Bermuda., *Deep-Sea Research Part II*, *48*(8 -9), 1471 – 1505, doi:10.1016/S0967-0645(00)00150-8.

De La Rocha, C. L., and U. Passow (2007), Factors influencing the sinking of POC and the efficiency of the biological carbon pump and the efficiency of the biological carbon pump, *Deep-Sea Research Part II*, *54*(5-7), 639–658, doi:10.1016/j.dsr2.2007.01.004.

De La Rocha, C. L., N. Nowald, and U. Passow (2008), Interactions between diatom aggregates, minerals, particulate organic carbon, and dissolved organic matter: Further implications for the ballast hypothesis, *Global Biogeochemical Cycles*, *22*(4), doi:10.1029/2007GB003156.

Denman, K., et al. (2007), Couplings between changes in the climate system and biogeochemistry, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Cambridge University Press.

Deutsch, C., J. L. Sarmiento, D. M. Sigman, N. Gruber, and J. P. Dunne (2007), Spatial coupling of nitrogen inputs and losses in the ocean, *Nature*, *445*(7124), 163–167.

DeVries, T., F. Primeau, and C. Deutsch (2012a), The sequestration efficiency of the biological pump, *Geophysical Research Letters*, *39*(13), doi:10.1029/2012GL051963, l13601.

DeVries, T., C. Deutsch, F. Primeau, B. Chang, and A. Devol (2012b), Global rates of water-column denitrification derived from nitrogen gas measurements, *Nature Geoscience*, *5*, 547–550.

DeVries, T., J.-H. Liang, and C. Deutsch (2014), A mechanistic particle flux model applied to the oceanic phosphorus cycle, *Biogeosciences*, *11*(19), 5381–5398, doi:10.5194/bg-11-5381-2014.

Dietze, H., and U. Loeptien (2013), Revisiting "nutrient trapping" in global coupled biogeochemical ocean circulation models, *Global Biogeochemical Cycles*, *27*(2), 265–284, doi:10.1002/gbc.20029.

Dietze, H., and A. Oschlies (2005), Modeling abiotic production of apparent oxygen utilisation in the oligotrophic subtropical North Atlantic, *Ocean Dynamics*, *55*(1), 28–33.

Dunne, J. P., R. A. Armstrong, A. Gnanadesikan, and J. L. Sarmiento (2005), Empirical and mechanistic models for the particle export ratio, *Global Biogeochemical Cycles*, *19*(4), GB4026, doi:10.1029/2004GB002390.

Dunne, J. P., J. L. Sarmiento, and A. Gnanadesikan (2007), A synthesis of global particle export from the surface ocean and cycling through the ocean interior and on the seafloor, *Global Biogeochemical Cycles*, *21*(4), GB4006, doi:10.1029/2006GB002907.

Dunne, J. P., B. Hales, and J. R. Toggweiler (2012), Global calcite cycling constrained by sediment preservation controls, *Global Biogeochemical Cycles*, *26*(3), GB3023, doi:10.1029/2010GB003935.

Duteil, O., et al. (2012), Preformed and regenerated phosphate in ocean general circulation models: can right total concentrations be wrong?, *Biogeosciences*, *9*(5), 1797–1807, doi:10.5194/bg-9-1797-2012.

Duteil, O., W. Koeve, A. Oschlies, D. Bianchi, E. Galbraith, I. Kriest, and R. Matear (2013), A novel estimate of ocean oxygen utilisation points to a reduced rate of respiration in the ocean interior, *Biogeosciences*, *10*(11), 7723–7738, doi:10.5194/bg-10-7723-2013.

Dutkiewicz, S., M. J. Follows, and J. G. Bragg (2009), Modeling the coupling of ocean ecology and biogeochemistry, *Global Biogeochemical Cycles*, *23*(4), doi:10.1029/2008GB003405.

Edwards, N., and R. Marsh (2005), Uncertainties due to transport-parameter sensitivity in an efficient 3-D ocean-climate model, *Climate Dynamics*, *24*(4), 415–433.

Edwards, N., and J. Shepherd (2002), Bifurcations of the thermohaline circulation in a simplified three-dimensional model of the world ocean and the effects of inter-basin connectivity, *Climate Dynamics*, *19*(1), 31–42, doi:10.1007/s00382-001-0207-7.

Engel, A., J. Szlosek, L. Abramson, Z. F. Liu, and C. Lee (2009a), Investigating the effect of ballasting by $CaCO_3$ in Emiliania huxleyi: I. formation, settling velocities and physical properties of aggregates, *Deep-Sea Research Part II*, *56*(18), 1396–1407, doi:10.1016/j.dsr2.2008.11.027.

Engel, A., L. Abramson, J. Szlosek, Z. Liu, G. Stewart, D.Hirschberg, and C. Lee (2009b), Investigating the effect of ballasting by $CaCO_3$ in Emiliania huxleyi, ii: Decomposition of particulate organic matter, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*(18), 1408 – 1419, doi:http://dx.doi.org/10.1016/j.dsr2.2008.11.028.

Eppley, R. (1972), Temperature and phytoplankton growth in the sea., *Fishery Bulletin*, *70*(4), 1063 – 1085.

Fay, A. R., and G. A. McKinley (2013), Global trends in surface ocean pCO_2 from in situ data, *Global Biogeochemical Cycles*, *27*(2), 541–557, doi:10.1002/gbc.20051.

Fay, A. R., and G. A. McKinley (2014), Global open-ocean biomes: mean and temporal variability, *Earth System Science Data*, *6*(2), 273–284, doi:10.5194/essd-6-273-2014.

Feely, R., C. Sabine, R. Schlitzer, J. Bullister, S. Mecking, and D. Greely (2004), Oxygen utilization and organic carbon remineralization in the upper water column of the pacific ocean, *Journal of Oceanography*, *60*(1), 45–52.

Fine, R. A. (2011), Observations of CFCs and SF6 as Ocean Tracers, *Annual Review of Marine Science*, *3*(1), 173–195, doi:10.1146/annurev.marine.010908.163933.

Fischer, G., and G. Karakas (2009), Sinking rates and ballast composition of particles in the atlantic ocean: implications for the organic carbon fluxes to the deep ocean, *Biogeosciences*, *6*(1), 85–102.

Follows, M. J., S. Dutkiewicz, S. Grant, and S. W. Chisholm (2007), Emergent biogeography of microbial communities in a model ocean, *Science*, *315*(5820), 1843–1846, doi:10.1126/science.1138544.

Fotheringham, A. S., M. E. Charlton, and C. Brunsdon (1998), Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis, *Environment and Planning A*, *30*(11), 1905–1927, doi:10.1068/a301905.

Fotheringham, A. S., C. Brunsdon, and M. E. Charlton (2002), *Geographically Weighted Regression: the analysis of spatially varying relationships*, Wiley, Chichester.

Francois, R., S. Honjo, R. Krishfield, and S. Manganini (2002), Factors controlling the flux of organic carbon to the bathypelagic zone of the ocean, *Global Biogeochemical Cycles*, *16*(4), doi:10.1029/2001GB001722.

Ganachaud, A., and C. Wunsch (2000), Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data, *Nature*, *408*, 453–457, doi:10.1038/35044048.

Garcia, H., R. Locarnini, T. Boyer, J. Antonov, M. Zweng, O. Baranova, and D. Johnson (2010), World Ocean Atlas 2009, Volume 4: Nutrients (phosphate, nitrate, silicate)., in *NOAA Atlas NESDIS 71*, edited by S. Levitus, U.S. Government Printing Office, Washington, D.C.

Gebbie, G., and P. Huybers (2010), Total Matrix Intercomparison: A method for determining the geometry of water-mass pathways, *Journal of Physical Oceanography*, *40*(8), 1710–1728, doi:10.1175/2010JPO4272.1.

Gehlen, M., L. Bopp, N. Ernprin, O. Aumont, C. Heinze, and O. Raguencau (2006), Reconciling surface ocean productivity, export fluxes and sediment composition in a global biogeochemical ocean model, *Biogeosciences*, *3*(4), 521–537.

Gloor, M., N. Gruber, T. M. C. Hughes, and J. L. Sarmiento (2001), Estimating net air-sea fluxes from ocean bulk data: Methodology and application to the heat cycle, *Global Biogeochemical Cycles*, *15*(4), 767–782, doi:10.1029/2000GB001301.

Gregoire, L., P. Valdes, A. Payne, and R. Kahana (2011), Optimal tuning of a GCM using modern and glacial constraints, *Climate Dynamics*, *37*(3-4), 705–719.

Gruber, N., and J. L. Sarmiento (2002), Large-scale biogeochemical-physical interactions in elemental cycles, in *The Sea*, edited by A. R. e. a. Robinson, John Wiley and Sons.

Gurney, K., et al. (2003), TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, *Tellus B*, *55*(2), 555–579, doi:10.1034/j.1600-0889.2003.00049.x.

Hain, M. P., and D. M. Sigman (2014), The biological pump in the past, in *Treatise on Geochemistry*, vol. 8, Elsevier.

Hansell, D., C. Carlson, D. Repeta, and R. Schlitzer (2009), Dissolved organic matter in the ocean: A controversy stimulates new insights, *Oceanography*, *22*(4), 202 – 211.

Hargreaves, J., J. Annan, N. Edwards, and R. Marsh (2004), An efficient climate forecasting method using an intermediate complexity Earth System Model and the ensemble Kalman filter, *Climate Dynamics*, *23*(7-8), 745–760, doi:10.1007/s00382-004-0471-4.

Harris, P., C. Brunsdon, and A. S. Fotheringham (2011), Links, comparisons and extensions of the geographically weighted regression model when used as a spatial predictor., *Stochastic Environmental Research and Risk Assessment*, *25*(2), 123 – 138, doi:10.1007/s00477-010-0444-6.

Heinze, C. (2004), Simulating oceanic $CaCO_3$ export production in the greenhouse, *Geophysical Research Letters*, *31*, doi:10.1029/2004GL020613.

Henson, S., R. Sanders, and E. Madsen (2012), Global patterns in efficiency of particulate organic carbon export and transfer to the deep ocean, *Global Biogeochemical Cycles*, *26*, GB1028, doi:10.1029/2011GB004099.

Henson, S. A., J. L. Sarmiento, J. P. Dunne, L. Bopp, I. Lima, S. C. Doney, J. John, and C. Beaulieu (2010), Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity, *Biogeosciences*, *7*(2), 621–640, doi:10.5194/bg-7-621-2010.

Henson, S. A., R. Sanders, E. Madsen, P. J. Morris, F. Le Moigne, and G. D. Quartly (2011), A reduced estimate of the strength of the ocean's biological carbon pump, *Geophysical Research Letters*, *38*(4), L04,606, doi:10.1029/2011GL046735.

Hofmann, M., and H. J. Schellnhuber (2009), Oceanic acidification affects marine carbon pump and triggers extended marine oxygen holes, *Proceedings of the National Academy of Sciences of hte United States of America*, *106*(9), 3017–3022, doi:10.1073/pnas.0813384106.

Holden, P. B., N. R. Edwards, S. A. Müller, K. I. C. Oliver, R. M. Death, and A. Ridgwell (2013), Controls on the spatial distribution of oceanic $\delta^{13}C_{DIC}$, *Biogeosciences*, *10*(3), 1815–1833, doi:10.5194/bg-10-1815-2013.

Honda, M. C., and S. Watanabe (2010), Importance of biogenic opal as ballast of particulate organic carbon (POC) transport and existence of mineral ballast-associated and residual POC in the Western Pacific Subarctic Gyre, *Geophysical Research Letters*, *37*, doi:10.1029/2009GL041521.

Hönisch, B., et al. (2012), The geological record of ocean acidification, *Science*, *335*(6072), 1058 – 1063, doi:10.1126/science.1208277.

Honjo, S., S. J. Manganini, R. A. Krishfield, and R. Francois (2008), Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: A synthesis of global sediment trap programs since 1983, *Progress in Oceanography*, *76*(3), 217–285, doi:10.1016/j.pocean.2007.11.003.

Honjo, S., et al. (2014), Understanding the role of the biological pump in the global carbon cycle: An imperative for ocean science., *Oceanography*, *27*(3), 10 – 16.

Hope, A. C. A. (1968), A simplified monte carlo significance test procedure, *Journal of the Royal Statistical Society, Series B*, *30*(3), 582–598.

Howard, M. T., A. M. E. Winguth, C. Klaas, and E. Maier-Reimer (2006), Sensitivity of ocean carbon tracer distributions to particulate organic flux parameterizations, *Global Biogeochemical Cycles*, *20*(3), doi:10.1029/2005GB002499.

Ito, T., and M. J. Follows (2005), Preformed phosphate, soft-tissue pump and atmospheric CO_2, *Journal of Marine Research*, *64*(4), 813 – 839.

Ito, T., M. J. Follows, and E. A. Boyle (2004), Is AOU a good measure of respiration in the oceans?, *Geophysical Research Letters*, *31*(17), L17,305, doi:10.1029/2004GL020900.

Iversen, M. H., and H. Ploug (2010), Ballast minerals and the sinking carbon flux in the ocean: carbon-specific respiration rates and sinking velocity of marine snow aggregates, *Biogeosciences*, *7*(9), 2613–2624, doi:10.5194/bg-7-2613-2010.

Jenkins, W. J. (1982), Oxygen utilization rates in north atlantic subtropical gyre and primary production in oligotrophic systems, *Nature*, *300*(5889), 246–248.

Jin, X., N. Gruber, J. P. Dunne, J. L. Sarmiento, and R. A. Armstrong (2006), Diagnosing the contribution of phytoplankton functional groups to the production and export of particulate organic carbon, $CaCO_3$, and opal from global nutrient and alkalinity distributions, *Global Biogeochemical Cycles*, *20*(2), GB2015, doi:10.1029/2005GB002532.

John, E., J. Wilson, P. Pearson, and A. Ridgwell (2014), Temperature-dependent remineralization and carbon cycling in the warm Eocene oceans, *Palaeogeography, Palaeoclimatology, Palaeoecology*, *413*(0), 158 – 166, doi:http://dx.doi.org/10.1016/j.palaeo.2014.05.019, selected Papers, Geologic Problem Solving with Microfossils3.

John, E. H., P. N. Pearson, H. K. Coxall, H. Birch, B. S. Wade, and G. L. Foster (2013), Warm ocean processes and carbon cycling in the Eocene, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *371*(2001), doi:10.1098/rsta.2013.0099.

Jouzel, J., et al. (2007), Orbital and millennial antarctic climate variability over the past 800,000 years, *Science*, *317*(5839), 793–796, doi:10.1126/science.1141038.

Key, R. M., et al. (2004), A global ocean carbon climatology: Results from global data analysis project (GLODAP), *Global Biogeochemical Cycles*, *18*(4), doi:10.1029/2004GB002247, gB4031.

Khatiwala, S. (2007), A computational framework for simulation of biogeochemical tracers in the ocean, *Global Biogeochemical Cycles*, *21*(3), GB3001, doi:10.1029/2007GB002923.

Khatiwala, S., M. Visbeck, and M. A. Cane (2005), Accelerated simulation of passive tracers in ocean circulation models, *Ocean Modelling*, *9*(1), 51 – 69, doi:http://dx.doi.org/10.1016/j.ocemod.2004.04.002.

Khatiwala, S., F. Primeau, and M. Holze (2012), Ventilation of the deep ocean constrained with tracer observations and implications for radiocarbon estimates of ideal mean age, *Earth and Planetary Science Letters*, *325-326*(1), 116 – 125, doi:doi:10.1016/j.epsl.2012.01.038.

Klaas, C., and D. E. Archer (2002), Association of sinking organic matter with various types of mineral ballast in the deep sea: Implications for the rain ratio, *Global Biogeochemical Cycles*, *16*(4), 1116, doi:10.1029/2001GB001765.

Kriest, I., and A. Oschlies (2008), On the treatment of particulate organic matter sinking in large-scale models of marine biogeochemical cycles, *Biogeosciences*, *5*(1), 55–72, doi:10.5194/bg-5-55-2008.

Kriest, I., and A. Oschlies (2011), Numerical effects on organic-matter sedimentation and remineralization in biogeochemical ocean models, *Ocean Modelling*, *39*(3–4), 275 – 283, doi:http://dx.doi.org/10.1016/j.ocemod.2011.05.001.

Kriest, I., and A. Oschlies (2013), Swept under the carpet: organic matter burial decreases global ocean biogeochemical model sensitivity to remineralization length scale, *Biogeosciences*, *10*(12), 8401–8422, doi:10.5194/bg-10-8401-2013.

Kriest, I., S. Khatiwala, and A. Oschlies (2010), Towards an assessment of simple global marine biogeochemical models of different complexity, *Progress in Oceanography*, *86*(3–4), 337 – 360, doi:http://dx.doi.org/10.1016/j.pocean.2010.05.002.

Kriest, I., A. Oschlies, and S. Khatiwala (2012), Sensitivity analysis of simple global marine biogeochemical models, *Global Biogeochemical Cycles*, *26*(2), GB2029, doi:10.1029/2011GB004072.

Krzywinski, M., and N. Altman (2014), Points of significance: Visualizing samples with box plots, *Nature Methods*, *11*(2), 119–120.

Kumar, S., and R. Lal (2011), Mapping the organic carbon stocks of surface soils using spatial interpolator, *Journal of Environmental Monitoring*, *13*, 3128 – 3135, doi:10.1039/c1em10520e.

Kwon, E. Y., and F. Primeau (2006), Optimization and sensitivity study of a biogeochemistry ocean model using an implicit solver and in situ phosphate data, *Global Biogeochemical Cycles*, *20*(4), GB4009, doi:10.1029/2005GB002631.

Kwon, E. Y., F. Primeau, and J. L. Sarmiento (2009), The impact of remineralization depth on the air-sea carbon balance, *Nature Geoscience*, *2*(9), 630–635.

Lam, P. J., and J. K. B. Bishop (2007), High biomass, low export regimes in the southern ocean, *Deep-Sea Research Part II*, *54*(5-7), 601–638, doi:10.1016/j.dsr2.2007.01.013.

Lam, P. J., S. C. Doney, and J. K. B. Bishop (2011), The dynamic ocean biological pump: Insights from a global compilation of particulate organic carbon, $CaCO_3$, and opal concentration profiles from the mesopelagic, *Global Biogeochemical Cycles*, *25*, GB3009, doi:10.1029/2010GB003868.

Lam, P. J., D. C. Ohnemus, and M. E. Auro (In Press), Size-fractionated major particle composition and concentrations from the US GEOTRACES north Atlantic zonal transect, *Deep Sea Research Part II: Topical Studies in Oceanography*, doi:http://dx.doi.org/10.1016/j.dsr2.2014.11.020.

Laws, E. A., P. G. Falkowski, W. O. Smith, H. Ducklow, and J. J. McCarthy (2000), Temperature effects on export production in the open ocean, *Global Biogeochemical Cycles*, *14*(4), 1231–1246.

Le Moigne, F. A. C., R. J. Sanders, M. Villa-Alfageme, A. P. Martin, K. Pabortsava, H. Planquette, P. J. Morris, and S. J. Thomalla (2012), On the proportion of ballast versus non-ballast associated carbon export in the surface ocean, *Geophysical Research Letters*, *39*(15), doi:10.1029/2012GL052980.

Le Moigne, F. A. C., S. A. Henson, R. J. Sanders, and E. Madsen (2013), Global database of surface ocean particulate organic carbon export fluxes diagnosed from the [234]Th technique, *Earth System Science Data*, *5*(2), 295–304, doi:10.5194/essd-5-295-2013.

Le Moigne, F. A. C., K. Pabortsava, C. L. J. Marcinko, P. Martin, and R. J. Sanders (2014), Where is mineral ballast important for surface export of particulate organic carbon in the ocean?, *Geophysical Research Letters*, *41*, doi:10.1002/2014GL061678.

Lee, C., et al. (2009), Particulate organic matter and ballast fluxes measured using time-series and settling velocity sediment traps in the northwestern Mediterranean Sea, *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*(18), 1420 – 1436, doi:http://dx.doi.org/10.1016/j.dsr2.2008.11.029.

Lenton, T., M. Williamson, N. Edwards, R. Marsh, A. Price, A. Ridgwell, J. Shepherd, S. Cox, and T. G. team. (2006), Millennial timescale carbon cycle and climate change in an efficient Earth system model, *Climate Dynamics*, *26*, doi:10.1007/s00382-006-0109-9),pages={687--711},url={http://eprints.soton.ac.uk/26700/}.

Lima, I. D., P. J. Lam, and S. C. Doney (2014), Dynamics of particulate organic carbon flux in a global ocean model, *Biogeosciences*, *11*(4), 1177–1198, doi:10.5194/bg-11-1177-2014.

Loeppky, J. L., J. Sacks, and W. J. Welch (2009), Choosing the sample size of a computer experiment: A practical guide, *Technometrics*, *51*(4), 366–376, doi:10.1198/TECH.2009.08040.

Longhurst, A. (1998), *Ecological Geography of the Sea*, Academic Press, San Diego.

Loubere, P., S. A. Siedlecki, and L. I. Bradtmiller (2007), Organic carbon and carbonate fluxes: Links to climate change, *Deep-Sea Research Part II*, *54*(5-7), 437–446, doi:10.1016/j.dsr2.2007.02.001.

Lutz, M., R. Dunbar, and K. Caldeira (2002), Regional variability in the vertical flux of particulate organic carbon in the ocean interior, *Global Biogeochemical Cycles*, *16*(3).

Lutz, M., K. Caldeira, R. Dunbar, and M. Behrenfeld (2007), 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: Oceans*, *112*(C10), C10,011, doi:10.1029/2006JC003706.

Mahowald, N. M., A. R. Baker, G. Bergametti, N. Brooks, R. A. Duce, T. D. Jickells, N. Kubilay, J. M. Prospero, and I. Tegen (2005), Atmospheric global dust cycle and iron inputs to the ocean, *Global Biogeochemical Cycles*, *19*(4), GB4025, doi:10.1029/2004GB002402.

Marinov, I., A. Gnanadesikan, J. L. Sarmiento, J. R. Toggweiler, M. Follows, and B. K. Mignone (2008), Impact of oceanic circulation on biological carbon storage in the ocean and atmospheric pCO_2, *Global Biogeochemical Cycles*, *22*(3), GB3007, doi:10.1029/2007GB002958.

Marsay, C., R. Sanders, S. Henson, K. Pabortsava, E. Achterberg, and R. Lampitt (2015), Attenuation of sinking particulate organic carbon flux through the mesopelagic ocean, *Proceedings of the National Academy of Sciences*, *112*(4), 1089–1094, doi:10.1073/pnas.1415311112.

Marsh, R., S. A. Müller, A. Yool, and N. R. Edwards (2011), Incorporation of the C-GOLDSTEIN efficient climate model into the GENIE framework: "eb_go_g" configurations of GENIE, *Geoscientific Model Development*, *4*(4), 957–992, doi: 10.5194/gmd-4-957-2011.

Martin, J. (1990), Glacial-interglacial CO_2 change: The Iron Hypothesis, *Paleoceanography*, *5*, doi:10.1029/PA005i001p00001.

Martin, J., G. Knauer, D. Karl, and W. Broenkow (1987), Vertex: carbon cycling in the northeast Pacific, *Deep Sea Research*, *43*(2), 267 – 285.

Matsumoto, K. (2007), Biology-mediated temperature control on atmospheric pCO_2 and ocean biogeochemistry, *Geophysical Research Letters*, *34*(20), doi:10.1029/2007GL031301.

Matsumoto, K., J. Sarmiento, and M. Brzezinski (2002), Silicic acid leakage from the southern ocean as a possible mechanism for explaining glacial atmospheric pCO_2, *Global Biogeochemical Cycles*, *16*, GB001,442, doi:10.1029/2001.

Matsumoto, K., T. Hashioka, and Y. Yamanaka (2007), Effect of temperature-dependent organic carbon decay on atmospheric pCO_2, *Journal of Geophysical Research: Biogeosciences*, *112*(G2), doi:10.1029/2006JG000187.

Matsumoto, K., K. S. Tokos, A. R. Price, and S. J. Cox (2008), First description of the Minnesota Earth System Model for Ocean biogeochemistry (MESMO 1.0), *Geoscientific Model Development*, *1*(1), 1–15, doi:10.5194/gmd-1-1-2008.

Mikaloff Fletcher, S. E., et al. (2006), Inverse estimates of anthropogenic CO_2 uptake, transport, and storage by the ocean, *Global Biogeochemical Cycles*, *20*(2), GB2002, doi:10.1029/2005GB002530.

Montgomery, D., E. Peck, and G. Vining (2006), *Introduction to Linear Regression Analysis*, Wiley, New Jersey.

Mortlock, R., and P. Froelich (1989), A simple method for the rapid determination of biogenic opal in pelagic marine sediments, *Deep Sea Research*, *36*(9), 1415–1426, doi:10.1016/0198-0149(89)90092-7.

Müller, S. A., F. Joos, N. R. Edwards, and T. F. Stocker (2006), Water mass distribution and ventilation time scales in a cost-efficient, three-dimensional ocean model, *Journal of Climate*, *19*(21), 5479–5499, doi:10.1175/JCLI3911.1.

Najjar, R. G., et al. (2007), Impact of circulation on export production, dissolved organic matter, and dissolved oxygen in the ocean: Results from phase II of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2), *Global Biogeochemical Cycles*, *21*(3), GB3007, doi:10.1029/2006GB002857.

Oka, A., S. Kato, and H. Hasumi (2008), Evaluating effect of ballast mineral on deep-ocean nutrient concentration by using an ocean general circulation model, *Global Biogeochemical Cycles*, *22*(3), doi:10.1029/2007GB003067.

Pace, M., G. Knauer, M. Karl, and J. Martin (1987), Primary production and vertical flux in the eastern Pacific Ocean., *Nature*, *325*, 803 – 804.

Passow, U. (2002), Transparent exopolymer particles (TEP) in aquatic environments, *Progress in Oceanography*, *55*(3–4), 287 – 333, doi:http://dx.doi.org/10.1016/S0079-6611(02)00138-6.

Passow, U. (2004), Switching perspectives: Do mineral fluxes determine particulate organic carbon fluxes or vice versa?, *Geochemistry Geophysics Geosystems*, *5*, doi:10.1029/2003GC000670.

Passow, U., and C. L. De la Rocha (2006), Accumulation of mineral ballast on organic aggregates, *Global Biogeochemical Cycles*, *20*(1), Q04,002, doi:10.1029/2005GB002579.

Paytan, A., and K. McLaughlin (2007), The Oceanic Phosphorus Cycle, *Chemical Reviews*, *107*(2), 563–576, doi:10.1021/cr0503613.

Ploug, H., M. H. Iversen, and G. Fischer (2008), Ballast, sinking velocity, and apparent diffusivity within marine snow and zooplankton fecal pellets: Implications for substrate turnover by attached bacteria, *Limnology and Oceanography*, *53*(5), 1878–1886, doi:10.4319/lo.2008.53.5.1878.

Primeau, F. (2006), On the variability of the exponent in the power law depth dependence of POC flux estimated from sediment traps, *Deep-Sea Research Part I-Oceanographic Research Papers*, *53*(8).

Ragueneau, O., S. Schultes, K. Bidle, P. Claquin, and B. La Moriceau (2006), Si and C interactions in the world ocean: Importance of ecological processes and implications for the role of diatoms in the biological pump, *Global Biogeochemical Cycles*, *20*(4), doi:10.1029/2006GB002688.

Redfield, A., B. Ketchum, and F. Richards (1963), The influence of organisms in the composition of sea water., in *The Sea*, edited by N. Hill, Wiley-Interscience: New York.

Ridgwell, A. (2001), Glacial-interglacial perturbations in the global carbon cycle., Ph.D. thesis, University of East Anglia.

Ridgwell, A., and S. Arndt (2014), Why dissolved organics matter: DOC in ancient oceans and past climate change, in *Biogeochemistry of Marine Dissolved Organic Matter*, edited by D. Hansell and C. Carlson, Elsevier.

Ridgwell, A., and J. C. Hargreaves (2007), Regulation of atmospheric CO_2 by deep-sea sediments in an Earth System model, *Global Biogeochemical Cycles*, *21*, GB2008, doi:doi:10.1029/2006GB002764.

Ridgwell, A., and R. E. Zeebe (2005), The role of the global carbonate cycle in the regulation and evolution of the earth system, *Earth and Planetary Science Letters*, *234*(3-4), 299–315.

Ridgwell, A., J. C. Hargreaves, N. R. Edwards, J. D. Annan, T. M. Lenton, R. Marsh, A. Yool, and A. Watson (2007a), Marine geochemical data assimilation in an efficient earth system model of global biogeochemical cycling, *Biogeosciences*, *4*(1), 87–104, doi:10.5194/bg-4-87-2007.

Ridgwell, A., I. Zondervan, J. C. Hargreaves, J. Bijma, and T. M. Lenton (2007b), Assessing the potential long-term increase of oceanic fossil fuel CO_2 uptake due to CO_2-calcification feedback, *Biogeosciences*, *4*(4), 481–492.

Ridgwell, A. J. (2003), An end to the "rain ratio" reign?, *Geochemistry Geophysics Geosystems*, *4*, doi:10.1029/2003GC000512.

Riebesell, U., A. Kortzinger, and A. Oschlies (2009), Sensitivities of marine carbon fluxes to ocean change, *Proceedings of the National Academy of Sciences of the United States of America*, *106*(49), 20,602–20,609, doi:10.1073/pnas.0813291106.

Riley, J., R. Sanders, F. Marsay, C.Moigne, E. Achterberg, and A. Poulton (2012), The relatice contribution of fast and slow sinking particles to ocean carbon export, *Global Biogeochemical Cycles*, *26*, doi:doi:10.1029/2011GB004085.

Roth, R., S. P. Ritz, and F. Joos (2014), Burial-nutrient feedbacks amplify the sensitivity of atmospheric carbon dioxide to changes in organic matter remineralisation, *Earth System Dynamics*, *5*(2), 321–343, doi:10.5194/esd-5-321-2014.

Sabine, C. L., et al. (2004), The oceanic sink for anthropogenic CO_2, *Science*, *305*(5682), 367–371, doi:10.1126/science.1097403.

Salter, I., A. Kemp, R. Lampitt, and M. Gledhill (2010), The association between biogenic and inorganic minerals and the amino acid composition of settling particles, *Limnology and Oceanography*, *55*, doi:10.4319/1o.2010.55.5.2207.

Sarmiento, J., N. Gruber, M. Brzezinski, and J. Dunne (2004a), High-latitude controls of thermocline nutrients and low latitude biological productivity, *Nature*, *427*, 56–60, doi:10.1038/nature02127.

Sarmiento, J. L., and N. Gruber (2006), *Ocean Biogeochemical Dynamics*, Princeton University Press.

Sarmiento, J. L., and J. R. Toggweiler (1984), A new model for the role of the oceans in determining atmospheric pCO$_2$, *Nature*, *308*(5960), 621–624.

Sarmiento, J. L., J. Dunne, A. Gnanadesikan, R. M. Key, K. Matsumoto, and R. Slater (2002), A new estimate of the $CaCO_3$ to organic carbon export ratio, *Global Biogeochemical Cycles*, *16*(4), 54–1–54–12, doi:10.1029/2002GB001919.

Sarmiento, J. L., et al. (2004b), Response of ocean ecosystems to climate warming, *Global Biogeochemical Cycles*, *18*(3), GB3003, doi:10.1029/2003GB002134.

Schlitzer, R. (2002a), Carbon export fluxes in the southern ocean: results from inverse modeling and comparison with satellite-based estimates, *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*(9–10), 1623 – 1644, doi:http://dx.doi.org/10.1016/S0967-0645(02)00004-8.

Schlitzer, R. (2002b), Interactive analysis and visualization of geoscience data with Ocean Data View, *Computers & Geosciences*, *28*(10), 1211 – 1218, doi:http://dx.doi.org/10.1016/S0098-3004(02)00040-7, shareware and freeware in the Geosciences II. A special issue in honour of John Butler.

Schlitzer, R. (2004), Export production in the equatorial and north Pacific derived from dissolved oxygen, nutrient and carbon data, *Journal of Oceanography*, *60*(1), 53 – 62.

Schmittner, A., N. Gruber, A. C. Mix, R. M. Key, A. Tagliabue, and T. K. Westberry (2013), Biology and air-sea gas exchange controls on the distribution of carbon isotope ratios (δ^{13}C) in the ocean, *Biogeosciences*, *10*(9), 5793–5816, doi:10.5194/bg-10-5793-2013.

Scholten, J. C., et al. (2001), Trapping efficiencies of sediment traps from the deep Eastern North Atlantic: the ^{230}Th calibration, *Deep-Sea Research Part II*, *48*(10), 2383–2408, doi:10.1016/S0967-0645(00)00176-4.

Segschneider, J., and J. Bendtsen (2013), Temperature-dependent remineralization in a warming ocean increases surface pCO_2 through changes in marine ecosystem composition, *Global Biogeochemical Cycles*, *27*(4), 1214–1225, doi:10.1002/2013GB004684.

Seiter, K., C. Hensen, and M. Zabel (2005), Benthic carbon mineralization on a global scale, *Global Biogeochemical Cycles*, *19*(1), doi:10.1029/2004GB002225.

Sigman, D. M., and E. A. Boyle (2000), Glacial/interglacial variations in atmospheric carbon dioxide, *Nature*, *407*(6806), 859–869.

Sigman, D. M., M. P. Hain, and G. H. Haug (2010), The polar ocean and glacial cycles in atmospheric CO_2 concentration, *Nature*, *466*(7302), 47–55.

Sonnerup, R. E., S. Mecking, and J. L. Bullister (2013), Transit time distributions and oxygen utilization rates in the northeast pacific ocean from chlorofluorocarbons and sulfur hexafluoride, *Deep Sea Research Part I: Oceanographic Research Papers*, *72*(0), 61 – 71, doi:http://dx.doi.org/10.1016/j.dsr.2012.10.013.

Stammer, D., K. Ueyoshi, A. Köhl, W. G. Large, S. A. Josey, and C. Wunsch (2004), Estimating air-sea fluxes of heat, freshwater, and momentum through global ocean data assimilation, *Journal of Geophysical Research: Oceans*, *109*, C05,023, doi:10.1029/2003JC002082.

Stanley, R. H. R., S. C. Doney, W. J. Jenkins, and D. E. Lott III (2012), Apparent oxygen utilization rates calculated from tritium and helium-3 profiles at the bermuda atlantic time-series study site, *Biogeosciences*, *9*(6), 1969–1983, doi:10.5194/bg-9-1969-2012.

Stemmann, L., G. Jackson, and G. Gorsky (2004), A vertical model of particle size distrubutions and fluxes in the midwater column that includes biological and physical process - Part II: application to a three year survey in the NW Mediterranean Sea, *Deep-Sea Research Part I*, *51*, doi:10.1016/j.dsr.2004.03.002.

Suess, E. (1980), Particle organic carbon flux in the oceans - surface productivity and oxygen utilization., *Nature*, *288*, 260 – 263.

Taucher, J., L. T. Bach, U. Riebesell, and A. Oschlies (2014), The viscosity effect on marine particle flux: A climate relevant feedback mechanism, *Global Biogeochemical Cycles*, *28*(4), 415–422, doi:10.1002/2013GB004728.

Taylor, K. (2001), Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183–7192, doi:10.1029/2000JD900719.

Teng, Y.-C., F. Primeau, J. Moore, M. Lomas, and A. Martiny (2014), Global-scale variations of the ratios of carbon to phosphorus in exported marine organic matter, *Nature Geoscience*, *7*, 895 – 898, doi:10.1038/ngeo2303.

Ternon, E., C. Guieu, M.-D. Loÿe-Pilot, N. Leblond, E. Bosc, B. Gasser, J.-C. Miquel, and J. Martín (2010), The impact of Saharan dust on the particulate export in the water column of the North Western Mediterranean Sea, *Biogeosciences*, *7*(3), 809–826, doi:10.5194/bg-7-809-2010.

Thunell, R., C. Benitez-Nelson, R. Varela, Y. Astor, and F. Muller-Karger (2007), Particulate organic carbon fluxes along upwelling-dominated continental margins: Rates and mechanisms, *Global Biogeochemical Cycles*, *21*(1), doi:10.1029/2006GB002793.

Urban, N., and T. Fricker (2010), A comparison of Latin hypercube and grid ensemble designs for the multivariate emulation of an Earth system model, *Computers & Geosciences*, *36*(6), 746 – 755, doi:http://dx.doi.org/10.1016/j.cageo.2009.11.004.

Usbeck, R. (1999), Modeling of marine biogeochemical cycles with an emphasis on vertical particles fluxes., Ph.D. thesis, Alfred-Wegener-Institut for Polar-und Meeresforschung.

USGS (2004), http://pubs.usgs.gov/of/2003/of03-001/index.htm.

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

WHOI (2007), http://www.whoi.edu/page.do?cid=10286pid=10979tid=282.

Williams, J. H. T., R. S. Smith, P. J. Valdes, B. B. B. Booth, and A. Osprey (2013), Optimising the FAMOUS climate model: inclusion of global carbon cycling, *Geoscientific Model Development*, *6*(1), 141–160, doi:10.5194/gmd-6-141-2013.

Williams, R., and M. J. Follows (2011), *Ocean Dynamics and the Carbon Cycle*, Cambridge: Cambridge University Press.

Wilson, J. D., S. Barker, and A. Ridgwell (2012), Assessment of the spatial variability in particulate organic matter and mineral sinking fluxes in the ocean interior: Implications for the ballast hypothesis, *Global Biogeochemical Cycles*, *26*(4), GB4011, doi:10.1029/2012GB004398.

Wong, C., F. Whitney, K. Crawford, K. Iseki, R. Matear, W. Johnson, J. Page, and D. Timothy (1999), Seasonal and interannual variability in particle fluxes of carbon, nitrogen and silicon from time series of sediment traps at Ocean Station P., *Deep-Sea Research Part II*, *46*, 2735 – 2760, doi:10.1016/S0967-0645(99)00082-X.

Yamanaka, Y., and E. Tajika (1996), The role of the vertical fluxes of particulate organic matter and calcite in the oceanic carbon cycle: Studies using an ocean biogeochemical general circulation model, *Global Biogeochemical Cycles*, *10*(2), 361–382, doi:10.1029/96GB00634.

Yao, X., and R. Schlitzer (2013), Assimilating water column and satellite data for marine export production estimation, *Geoscientific Model Development*, *6*(5), 1575–1590, doi:10.5194/gmd-6-1575-2013.

Yool, A., E. E. Popova, and T. R. Anderson (2013), MEDUSA-2.0: an intermediate complexity biogeochemical model of the marine carbon cycle for climate change and ocean acidification studies, *Geoscientific Model Development*, *6*(5), 1767–1811, doi: 10.5194/gmd-6-1767-2013.

Yu, E. F., R. Francois, M. P. Bacon, S. Honjo, A. P. Fleer, S. J. Manganini, M. M. R. van der Loeff, and V. Ittekot (2001), Trapping efficiency of bottom-tethered sediment traps estimated from the intercepted fluxes of Th-230 and Pa-231, *Deep-Sea Research Part I*, *48*(3), 865–889, doi:10.1016/S0967-0637(00)00067-4.

Zeebe, R. E. (2012), Loscar: Long-term ocean-atmosphere-sediment carbon cycle reservoir model v2.0.4, *Geoscientific Model Development*, *5*(1), 149–166, doi:10.5194/gmd-5-149-2012.

Zeebe, R. E., and A. Ridgwell (2011), Past changes in ocean carbonate chemistry, in *Ocean Acidification*, edited by J.-P. Gattuso and L. Hansson.

Zeebe, R. E., and D. Wolf-Gladrow (2001), *CO_2 in Seawater: Equilibrium, Kinetics, Isotopes*, Elsevier.

Ziegler, M., P. Diz, I. Hall, and R. Zahn (2013), Millennial-scale changes in atmospheric CO_2 levels linked to the Southern Ocean carbon isotope gradient and dust flux, *Nature Geoscience*, *6*, 457–461, doi:10.1038/ngeo1782.

Ziveri, P., B. de Bernardi, K. H. Baumann, H. M. Stoll, and P. G. Mortyn (2007), Sinking of coccolith carbonate and potential contribution to organic carbon ballasting in the deep ocean, *Deep-Sea Research Part II*, *54*(5-7), 659–675, doi:10.1016/j.dsr2.2007.01.006.

APPENDIX A

Sediment Trap Data and GWR Output

A.1 Sediment Trap Data

The sediment trap dataset used in Chapter 3 is primarily comprised of data from the Joint Global Ocean Flux Study (JGOFS) available in its full form online at:

http://usjgofs.whoi.edu/mzweb/syndata.html#Sed_Traps.

Full methodological details and references can be found online through this link. Additional data was collected from the World Data Centre for Environmental Sciences (WDC-MARE) online database. Where data was missing in the JGOFS database, the original papers were consulted to amend details where appropriate. This is detailed in Table A1. A full reference list for data used is located after the table.

Flux data was reported as particulate inorganic carbon (PIC), biogenic silica (bSi). These were converted to fluxes of Calcium Carbonate ($CaCO_3$), Opal and Lithogenic material using the following relationships:

$$POM = POC * 1.86 \tag{A.1}$$

$$CaCO_3 = PIC * 8.33 \tag{A.2}$$

$$Opal = bSi * 2.14 \tag{A.3}$$

$$\begin{aligned}
\text{Litho.} &= \text{Total Mass} - (\text{POM} + \text{CaCO}_3 + \text{Opal}) \text{ if Litho.} > 0 \\
&= 0 \text{ otherwise}
\end{aligned} \tag{A.4}$$

Table A.1: Sediment trap data and local statistics calculated using Geographically Weighted Regression in Chapter 3. Data are arranged by latitude in each basin. Province refers to biogeochemical provinces as defined by *Longhurst* (1998). Latitude and longitude are in degrees. All flux data (Total, POC, PIC, bSi and Litho) are given in g m^{-2} year^{-1}. Depth is given in m. Latitude and Longitude are given in degrees North and East. All numbered references indicated are found below the table.

Province	Lat	Lon	Depth	Total	POC	PIC	bSi	Litho	β_0	β_{CaCO3}	β_{opal}	β_{Litho}	R^2	References
S. Ocean														
ANTA	-50.13	5.83	3196	29.4	4.61	0.9	4.22	4.3	0.097	0.121	0.042	-0.008	0.727	[15]; [14]; [50]
ANTA	-50.18	5.9	3110	6.44	0.51	0.3	1.02	0.81	0.098	0.121	0.042	-0.008	0.727	[15]; [14]; [50]
ANTA	-52	61.52	4000	41.49	0.57	0.6	14.32	4.79	0.286	0.112	-0.001	0.029	0.771	[53]
ANTA	-54.33	-3.38	2194	28.82	0.51	0.21	8.22	8.53	0.082	0.117	0.046	-0.010	0.723	[15]; [14]
ANTA	-54.33	-3.3	2251	3.07	0.63	0.08	0.22	0.76	0.082	0.117	0.046	-0.010	0.723	[15]; [14]
ANTA	-54.33	-3.3	2251	7.38	0.13	0.02	2.36	1.92	0.082	0.117	0.046	-0.010	0.723	[15]; [14]
ANTA	-56.9	-170.17	4224	24.55	0.73	1.43	3.85	3.04	0.168	0.069	0.008	0.029	0.688	[33]
ANTA	-57	-37	2000	8	0.1	0.12	3.29	0	0.080	0.111	0.047	-0.008	0.707	[25]
APLR	-62.08	-40.6	2453	3.36	0.22	0	0.78	1.28	0.118	0.111	0.035	0.000	0.704	[44]
ANTA	-62.26	-57.53	1588	107.3	2.32	0.62	18.05	59.19	0.160	0.102	0.024	0.010	0.696	[50]; [13] [a]
APLR	-63	71	3500	4.72	0.16	0.03	1.44	1.09	0.273	0.113	0.000	0.027	0.762	[53]
APLR	-63.15	-42.68	2966	1.35	0.09	0	0.32	0.5	0.129	0.110	0.032	0.003	0.703	[44]
Pacific														
BERS	58	179	3137	58.03	2.02	0.79	16.22	12.98	0.388	0.035	0.030	0.025	0.786	[26]; [54]

Continued on next page

Province	Lat	Lon	Depth	Total	POC	PIC	bSi	Litho	β_0	β_{CaCO3}	β_{opal}	β_{Litho}	R^2	References
BERS	53.5	-177	3193	40.22	1.93	0.8	12.55	3.11	0.432	0.026	0.033	0.023	0.783	[48]; [47]
BERS	53.5	-177	3200	70.55	2.34	1.1	25.02	3.49	0.432	0.026	0.033	0.023	0.783	[48]; [47]
BERS	53.5	-177	3200	62.74	2.56	0.99	19.7	7.57	0.432	0.026	0.033	0.023	0.783	[48]; [47]
BERS	53.5	-177	3199	100.59	3.43	1.75	31.3	12.65	0.432	0.026	0.033	0.023	0.783	[48]; [47]
BERS	53.5	-177	3198	54.79	2.19	0.99	14.04	12.42	0.432	0.026	0.033	0.023	0.783	[48]; [47]
PSAG	50	-145	3800	48.2	2.46	2.23	8.8	6.22	0.503	0.022	0.033	0.022	0.844	[52]
PSAG	50	-145	3800	30.4	0.82	1.72	5.18	3.46	0.503	0.022	0.033	0.022	0.844	[52]
PSAG	50	-145	3800	44.1	1.1	2.71	7.88	2.62	0.503	0.022	0.033	0.022	0.844	[52]
PSAG	50	-145	3800	30.7	0.83	1.45	6.94	2.23	0.503	0.022	0.033	0.022	0.844	[52]
PSAG	50	-145	3800	32.4	0.85	2.09	4.45	3.89	0.503	0.022	0.033	0.022	0.844	[52]
PSAG	50	-145	3800	22.1	1.29	1.25	3.76	1.24	0.503	0.022	0.033	0.022	0.844	[52]
PSAG	50	165	3260	42.52	0.99	1.39	11.68	4.1	0.287	0.040	0.032	0.024	0.769	[22]
PSAG	50	165	5090	31.76	0.88	1.06	7.62	4.99	0.287	0.040	0.032	0.024	0.769	[22]
PSAG	50	165	5090	34.64	0.8	0.84	10.11	4.52	0.287	0.040	0.032	0.024	0.769	[22]
PSAG	49	-174	4826	21.06	1.24	1.24	4.23	0	0.450	0.022	0.034	0.022	0.784	[48]
PSAG	49	-174	4806	33.91	1.61	0.99	8.66	4.14	0.450	0.022	0.034	0.022	0.784	[48]
PSAG	49	-174	4833	23.07	1.13	0.84	6.04	1.05	0.450	0.022	0.034	0.022	0.784	[48]
PSAG	49	-174	4822	42.34	2.08	1.64	10.49	2.36	0.450	0.022	0.034	0.022	0.784	[48]
PSAG	49	-174	4774	45.26	1.68	1.24	11.72	6.73	0.450	0.022	0.034	0.022	0.784	[48]
PSAG	44	155	2960	57.45	2.23	1.83	13.93	8.25	0.162	0.053	0.031	0.025	0.769	[22]
PSAG	44	155	4999	43.47	1.61	1.24	10.22	8.28	0.162	0.053	0.031	0.025	0.769	[22]
PSAG	44	155	4990	48.55	1.46	1.1	12.71	9.47	0.162	0.053	0.031	0.025	0.769	[22]
OCAL	42.19	-127.58	1500	60.5	2.2	1.09	7.77	30.7	0.458	0.028	0.034	0.021	0.881	[10]; [41] [b]
OCAL	42.09	-125.77	1500	190.1	4.9	1.79	22.3	118.35	0.461	0.028	0.034	0.021	0.883	[10]; [41] [b]
OCAL	41.55	-132	1500	10.3	0.9	0.5	1.54	1.17	0.441	0.030	0.034	0.021	0.875	[10]; [41] [b]
KURO	40	165	2986	35	1.31	1.61	7.39	3.34	0.181	0.048	0.033	0.022	0.768	[22]
KURO	40	165	5020	28.32	1.1	1.17	6.21	3.24	0.181	0.048	0.033	0.022	0.768	[22]

Province	Lat	Lon	Depth	Total	POC	PIC	bSi	Litho	β_0	β_{CaCO3}	β_{opal}	β_{Litho}	R^2	References
KURO	40	165	5020	40.66	1.35	1.28	11.29	3.33	0.181	0.048	0.033	0.022	0.768	[22]
CHIN	39.67	132.4	2800	96	2.94	1.88	19.51	33.12	0.013	0.075	0.027	0.030	0.781	[23]
OCAL	39.49	-127.69	3785	39.7	1.27	0.66	3.97	23.34	0.428	0.031	0.034	0.021	0.883	[10] [b]
KURO	34.42	177.74	2848	28.36	1.56	1.87	1.32	7.06	0.224	0.047	0.033	0.022	0.801	[37]; [38]
KURO	34.17	141.98	3423	35.77	0.79	1.9	3.8	10.34	0.005	0.074	0.028	0.030	0.781	[19]
KURO	34.17	141.98	5429	26.65	0.64	0.96	2.98	11.09	0.005	0.074	0.028	0.030	0.781	[19]
KURO	30	175	3873	17.3	0.97	1.32	0.64	3.13	0.174	0.052	0.033	0.022	0.807	[37]; [38]
KURO	27	127	1547	47.32	1.5	1.7	0.38	29.56	-0.082	0.096	0.022	0.030	0.825	[22]
WARM	12.01	134.29	4300	2.74	0.16	0.18	0.2	0.52	-0.092	0.096	0.023	0.032	0.840	[39]
PNEC	11	-140	1600	15.21	0.86	1.01	1.93	1.07	0.302	0.042	0.034	0.020	0.879	[8]; [10]
PNEC	11	-140	3400	13.19	0.59	0.87	1.81	0.97	0.302	0.042	0.034	0.020	0.879	[8]; [10]
PNEC	9	-139.98	2250	8.26	0.56	0.6	0.79	0.53	0.296	0.043	0.034	0.020	0.879	[31]
PNEC	9	-139.98	2150	8.65	0.55	0.66	0.74	0.55	0.296	0.043	0.034	0.020	0.879	[31]
PNEC	8.83	-103.98	3150	19.8	1.38	2.51	3.32	0	0.269	0.043	0.039	0.019	0.917	[10] [c]
PNEC	8.75	-104	2700	19.8	1.38	0.79	5.89	0	0.268	0.043	0.039	0.019	0.917	[9]
PNEC	6.57	-92.77	3565	25.5	0.89	0.79	1.81	13.39	0.232	0.044	0.044	0.017	0.932	[10] [c]
PNEC	5.02	-139.79	2100	27.34	1.64	2.15	2.42	1.2	0.286	0.043	0.034	0.020	0.881	[31]
PNEC	5.02	-139.79	3800	24.55	1.4	1.89	2.41	1.04	0.286	0.043	0.034	0.020	0.881	[31]
WARM	5.01	138.83	3130	4.08	0.25	0.32	0.29	0.33	-0.086	0.097	0.023	0.028	0.828	[39]
WARM	4.13	136.28	4574	34.64	2.21	1.42	5.35	7.25	-0.085	0.098	0.022	0.030	0.838	[38]; [36]
WARM	3	135	1590	58.66	3.41	3.13	6.22	12.93	-0.082	0.099	0.022	0.030	0.841	[38]; [36]
WARM	3	135	3900	56.68	2.86	2.81	6.19	14.71	-0.082	0.099	0.022	0.030	0.841	[38]; [36]
PNEC	2.01	-140.14	2203	26.76	1.28	2.16	2.83	0.33	0.277	0.044	0.035	0.020	0.881	[31]
PNEC	1.04	-138.94	4445	37.01	1.3	2.51	5	2.98	0.274	0.044	0.035	0.020	0.882	[10]
PNEC	1	-140	1895	26.91	1.55	1.82	4.3	0	0.275	0.044	0.035	0.020	0.882	[8]; [10]
PNEC	1	-140	3495	21.78	1.19	1.48	3.26	0.26	0.275	0.044	0.035	0.020	0.882	[8]; [10]
PNEC	1	-140	1883	54.39	2.47	3.56	9.57	0	0.275	0.044	0.035	0.020	0.882	[8]; [10]

Province	Lat	Lon	Depth	Total	POC	PIC	bSi	Litho	β_0	β_{CaCO3}	β_{opal}	β_{Litho}	R^2	References
PNEC	1	-140	2908	41.74	1.89	2.75	7.29	0	0.275	0.044	0.035	0.020	0.882	[8]; [10]
PEQD	0	-140	2284	35.17	1.67	2.8	3.14	2.02	0.272	0.044	0.035	0.020	0.882	[31]
PEQD	0	-140	3618	34.75	1.6	2.77	4.06	0.01	0.272	0.044	0.035	0.020	0.882	[31]
PEQD	-1.95	-139.75	3593	31.15	1.32	2.41	3.87	0.34	0.267	0.044	0.035	0.020	0.883	[31]
SPSG	-5	-140	2099	22.44	1	1.73	2.18	1.5	0.258	0.045	0.036	0.020	0.884	[31]
SPSG	-5	-140	2209	23.57	0.99	1.76	2.57	1.57	0.258	0.045	0.036	0.020	0.884	[31]
SPSG	-5	-140	2316	23.36	1.02	1.79	2.37	1.48	0.258	0.045	0.036	0.020	0.884	[31]
SPSG	-12	-135	3594	7.84	0.26	0.74	0.48	0.17	0.228	0.047	0.036	0.019	0.889	[31]
CHIL	-29.5	-73.2	2300	65.1	2.8	3.72	5.29	17.58	0.098	0.067	0.037	0.009	0.728	[20]

Atlantic

BPLR	81.07	138.9	1550	16.7	1.05	0.06	1.36	11.34	0.209	0.075	0.019	0.035	0.789	[11]
ARCT	78.86	1.37	2400	5	0.24	0.1	0.28	3.12	-0.087	0.090	0.018	0.064	0.810	[28]; [25]; [29]
ARCT	75.86	11.47	1700	24.63	1.69	0.63	0.87	14.38	-0.116	0.086	0.020	0.069	0.804	[28]; [25]; [29]
ARCT	74.58	-6.72	2823	9.89	0.4	0.38	1.22	3.37	0.025	0.081	0.018	0.055	0.797	[28]; [25]; [29]
ARCT	73.72	14.82	1692	41.25	1.91	1.72	2.06	18.96	-0.152	0.085	0.022	0.073	0.803	[40]
ARCT	72.5	-8	2200	9.44	0.36	0.33	0.42	5.12	0.061	0.076	0.018	0.053	0.791	[49]
ARCT	70	0	3000	34.43	3.28	2.02	0.81	9.77	-0.043	0.075	0.023	0.067	0.788	[49]
ARCT	70	0	3000	49.01	3.48	3.25	1.33	12.62	-0.043	0.075	0.023	0.067	0.788	[49]
ARCT	70	0	3000	60.32	3.93	1.74	0.58	37.27	-0.043	0.075	0.023	0.067	0.788	[49]
SARC	69.5	10	2760	19.95	0.81	1.16	0.51	7.69	-0.115	0.078	0.027	0.072	0.796	[28]; [25]
NADR	54.53	-21.07	2880	21.75	0.74	1.84	1.33	2.2	0.171	0.056	0.024	0.051	0.744	[46]
NADR	49.08	-13.42	3220	41.25	2.97	2	3.18	12.26	0.059	0.046	0.102	0.057	0.747	[46]; [2]; [1]
NADR	48.98	-13.75	4000	40.04	2.48	2.69	0.82	11.26	0.061	0.046	0.101	0.057	0.746	[46]; [2]; [1]
NADR	47.83	-19.5	3100	22.27	1.89	1.39	1.72	3.5	0.123	0.048	0.068	0.053	0.736	[35]

Province	Lat	Lon	Depth	Total	POC	PIC	bSi	Litho	β_0	β_{CaCO3}	β_{opal}	β_{Litho}	R^2	References
NADR	47.8	-19.78	3530	14.02	0.67	1.09	0.58	2.45	0.127	0.049	0.066	0.052	0.736	[46]
NADR	47.8	-19.8	2000	29.38	2.29	2.57	1.72	0.03	0.127	0.049	0.066	0.052	0.736	[46]
NADR	47.8	-19.8	3500	24.27	1.04	2.1	1.16	2.36	0.127	0.049	0.066	0.052	0.736	[46]
NADR	47.72	-20.87	3718	25.21	0.99	1.85	2.59	2.42	0.142	0.050	0.057	0.051	0.736	[30]; [27] [d]
NAST	33.82	-21.02	4564	21.38	0.93	1.63	0.98	3.98	0.024	0.053	0.119	0.040	0.761	[30]; [27] [d]
NAST	33.15	-21.98	4000	19.93	0.78	1.77	0.62	2.41	0.023	0.054	0.117	0.039	0.760	[46]
NAST	32.95	-22.02	2000	12.08	0.51	1.2	0.28	0.54	0.022	0.054	0.118	0.039	0.760	[46]
NAST	32.95	-22.02	4150	12.01	0.38	1.07	0.38	1.58	0.022	0.054	0.118	0.039	0.760	[46]
NAST	32.73	-70.82	2835	20	0.76	1.33	0.84	5.71	0.231	0.056	0.025	0.027	0.812	[10]; [21] [b]
NAST	32.08	-64.25	1500	14.13	0.88	0.98	3.47	0	0.210	0.057	0.018	0.029	0.789	[5]; [6]; [4] [e]
NAST	32.08	-64.25	3200	13.21	0.62	0.92	4.16	0	0.210	0.057	0.018	0.029	0.789	[5]; [6]; [4] [e]
NAST	31.83	-64.17	1500	11.4	0.88	0.79	0.77	1.53	0.208	0.057	0.018	0.029	0.788	[4] [7]; [3]
NAST	31.83	-64.17	3200	12.8	0.62	0.89	0.85	2.41	0.208	0.057	0.018	0.029	0.788	[4] [7]; [3]
NAST	29.12	-15.42	3075	18.07	0.83	1.02	0.33	7.32	-0.008	0.057	0.141	0.032	0.771	[46]; [43]
NAST	24.55	-22.83	3870	15.11	0.73	0.89	0.43	5.42	-0.033	0.062	0.134	0.026	0.767	[35]
NATR	21.3	-20.7	3568	31.53	1.56	1.94	0.68	11.01	-0.065	0.066	0.149	0.018	0.771	[16]
NATR	21.145	-20.687	3502	56.88	1.73	3.53	1.14	21.82	-0.066	0.066	0.150	0.018	0.771	[16]
NATR	21.138	-20.672	3557	48.05	2.05	2.8	0.91	18.97	-0.066	0.066	0.150	0.018	0.771	[16]
NAST	21.13	-20.67	2195	52.1	1.62	3.2	1.15	19.97	-0.066	0.066	0.150	0.018	0.771	[12]; [51] [d]
NAST	21.13	-20.67	3502	49.8	1.69	3	0.96	19.61	-0.066	0.066	0.150	0.018	0.771	[12]; [51] [d]
NATR	20.755	-19.742	2195	64.28	2.62	3.31	2.44	26.61	-0.072	0.067	0.155	0.017	0.772	[16]
NATR	11.5	-21.1	4435	11.24	0.45	0.28	0.4	7.21	-0.147	0.083	0.157	-0.006	0.769	[16]
WTRA	0	-23.4	3139	21.56	1.18	1.51	0.86	4.95	-0.197	0.108	0.128	-0.039	0.750	[16]
WTRA	0	-23.4	3173	25.03	1.31	2.03	1.05	3.44	-0.197	0.108	0.128	-0.039	0.750	[16]
WTRA	0	-23.4	3180	33.97	1.89	2.59	1.55	5.56	-0.197	0.108	0.128	-0.039	0.750	[16]
WTRA	0	-23.5	3170	26.21	1.11	2.24	0.75	3.88	-0.197	0.108	0.128	-0.039	0.750	[16]
WTRA	-3.9667	-25.65	4630	18.02	0.81	1.49	0.66	2.69	-0.189	0.110	0.124	-0.044	0.743	[16]

Province	Lat	Lon	Depth	Total	POC	PIC	bSi	Litho	β_0	β_{CaCO3}	β_{opal}	β_{Litho}	R^2	References
WTRA	-3.988	-25.583	4555	18.98	0.94	1.61	0.77	2.17	-0.189	0.111	0.125	-0.045	0.743	[16]
SATL	-7.472	-28.123	4410	9.3	0.47	0.82	0.16	1.25	-0.172	0.111	0.120	-0.045	0.737	[16]
SATL	-7.5	-28.2	4736	4.83	0.23	0.43	0.06	0.69	-0.172	0.111	0.120	-0.045	0.737	[16]
SATL	-7.5	-28.2	4705	8.38	0.5	0.65	0.27	1.46	-0.172	0.111	0.120	-0.045	0.737	[16]
SATL	-7.5	-28.1	4456	9.82	0.51	0.87	0.23	1.13	-0.172	0.111	0.120	-0.045	0.737	[16]
SATL	-7.5167	-28.0333	5031	6.71	0.38	0.53	0.16	1.25	-0.172	0.111	0.121	-0.046	0.737	[16]
SATL	-11.55	-28.52	4515	4.02	0.2	0.35	0.09	0.54	-0.162	0.112	0.120	-0.048	0.732	[16]
SATL	-20.04	9.1617	1647	24.63	2.86	1.54	1.34	3.61	-0.015	0.133	0.093	-0.066	0.745	[16]
SATL	-20.05	9.15	1648	36.93	3.77	2.87	1.41	2.99	-0.015	0.133	0.093	-0.066	0.745	[16]
SATL	-20.07	9.17	1640	30.5	3.32	2.29	1.35	2.36	-0.015	0.133	0.092	-0.065	0.745	[51]; [12]; [14] [d]
SATL	-20.07	9.17	1648	25.6	2.94	1.68	1.38	3.18	-0.015	0.133	0.092	-0.065	0.745	[51]; [12]; [14]
SATL	-20.07	9.17	1648	29.5	2.24	1.98	1.84	4.9	-0.015	0.133	0.092	-0.065	0.745	[51]; [12]; [14] [d]
NATR	-20.07	9.1667	1640	60.84	6.26	4.06	3.96	6.9	-0.015	0.133	0.092	-0.065	0.745	[16]
SATL	-28.9	14.6	2516	31.99	1.61	2.67	0.93	4.76	0.075	0.123	0.058	-0.029	0.750	[14] [f]

Indian

Province	Lat	Lon	Depth	Total	POC	PIC	bSi	Litho	β_0	β_{CaCO3}	β_{opal}	β_{Litho}	R^2	References
INDE	17.45	89.6	1727	51.81	2.7	1.29	3.55	28.45	-0.190	0.102	0.015	0.050	0.905	[34]; [45]
INDE	17.45	89.6	1727	52.22	2.99	1.97	5.04	19.46	-0.190	0.102	0.015	0.050	0.905	[34]; [45]
ARAB	17.4	58.8	3141	79.4	4.83	4.67	7.54	15.38	-0.206	0.096	0.029	0.061	0.905	[32]
ARAB	17.21	59.59	1857	83.04	5.96	4.72	7.38	16.84	-0.208	0.096	0.029	0.060	0.905	[32]
ARAB	17.21	59.59	2871	81.19	4.87	4.65	6.89	18.65	-0.208	0.096	0.029	0.060	0.905	[32]
ARAB	16	60	3020	51.5	2.6	3.38	4.88	8.07	-0.206	0.097	0.027	0.060	0.904	[17]; [18]; [42]
ARAB	16	60	3020	43.4	2.5	3.02	4.46	4.05	-0.206	0.097	0.027	0.060	0.904	[17]; [18]; [42]
ARAB	16	60	3020	53.2	3.3	3.91	4.08	5.76	-0.206	0.097	0.027	0.060	0.904	[17]; [18]; [42]
ARAB	16	60	3020	75.6	4.5	4.27	10.08	10.09	-0.206	0.097	0.027	0.060	0.904	[17]; [18]; [42]
MONS	15.23	89.17	1717	31.74	2.04	1.47	2.87	9.56	-0.194	0.102	0.015	0.051	0.906	[34]; [45]

Province	Lat	Lon	Depth	Total	POC	PIC	bSi	Litho	β_0	β_{CaCO3}	β_{opal}	β_{Litho}	R^2	References
MONS	15.23	89.17	1717	32.19	2.18	1.26	3.29	10.6	-0.194	0.102	0.015	0.051	0.906	[34]; [45]
MONS	15	68	2800	41.5	2.8	2.52	2.29	10.4	-0.211	0.101	0.018	0.057	0.903	[17]; [18]; [42]
MONS	15	68	2800	22.4	1.4	1.33	1.42	5.68	-0.211	0.101	0.018	0.057	0.903	[17]; [18]; [42]
MONS	15	68	2800	40.3	2.2	2.38	3.25	9.43	-0.211	0.101	0.018	0.057	0.903	[17]; [18]; [42]
MONS	15	68	2800	29.3	1.9	1.97	2.33	4.37	-0.211	0.101	0.018	0.057	0.903	[17]; [18]; [42]
MONS	14	64	2900	34.3	1.9	2.72	1.63	4.62	-0.208	0.100	0.020	0.059	0.902	[17]; [18]; [42]
MONS	14	64	2900	18.4	1.1	1.44	0.79	2.67	-0.208	0.100	0.020	0.059	0.902	[17]; [18]; [42]
MONS	14	64	2900	41.4	2.6	2.89	2.08	8.04	-0.208	0.100	0.020	0.059	0.902	[17]; [18]; [42]
MONS	13.15	84.37	2282	43.16	2.51	1.95	3.39	14.99	-0.206	0.102	0.015	0.053	0.908	[34]; [45]
MONS	13.15	84.37	2282	49.77	2.94	1.92	4.37	18.96	-0.206	0.102	0.015	0.053	0.908	[34]; [45]
MONS	13.15	84.37	2282	63.52	3.14	1.73	5.74	30.99	-0.206	0.102	0.015	0.053	0.908	[34]; [45]
MONS	10	65	2363	24.4	1.39	1.82	1.25	3.98	-0.190	0.102	0.017	0.058	0.897	[32]
MONS	10	65	3915	21.31	1.21	1.53	1.11	3.94	-0.190	0.102	0.017	0.058	0.897	[32]
MONS	4.46	87.3	3006	38.37	2.09	2.23	3	9.49	-0.189	0.103	0.015	0.051	0.907	[34]; [45]
MONS	4.46	87.3	3006	30.97	1.85	1.81	3.57	4.81	-0.189	0.103	0.015	0.051	0.907	[34]; [45]

[a] All flux values from [15]

[b] Depth from [54]

[c] Depth amended from within reference(s)

[d] Depth modified from JGOFS

[e] All fluxes amended from within reference(s)

[f] PIC and bSi values from [15]

References for Table A1:

[1] Antia, A., B. von Bodungen, and R. Peinert, (1999) Particle flux across the mid-European continental margin, Deep-Sea Research I, 46, 1999-2024.

[2] Antia, A., W. Koeve, G. Fischer, T. Blanz, D. Schulz-Bull, J. Scholten, S. Neuer, K. Kremling, J. Kuss, R. Peinert, D. Hebbeln, U. Bathmann, M. Conte, U. Fehner, and B. Zeitzschel, (2001) Basin-wide particulate carbon flux in the Atlantic Ocean: Regional export patterns and potential for atmospheric CO_2 sequestration, Global Biogeochemical Cycles, 15 (4), 845-862.

[3] Bacon, M.P., C.-A. Huh, A.P. Fleer, and W.G. Deuser, (1985) Seasonality in the flux of natural radionuclides and plutonium in the deep Sargasso Sea, Deep-Sea Research, 32 (3), 273-286.

[4] Conte, M., N. Ralf, and E.H. Ross, (2001) Seasonal and interannual variability in deep ocean particle fluxes at the Ocean Flux Program (OFP)/ Bermuda Atlantic Time-Series (BATS) site in the western Sargasso Sea near Bermuda, Deep-Sea Research II, 48, 1471-1505.

[5] Deuser, W. G., (1987) Variability of hydrography and particle flux: transient and long-term relationships. Particle Flux in the Ocean. E. T. Degens, E. Izdar and S. Honjo. Hamburg, Germany, SCOPE/UNEP. 62: 179-193.

[6] Deuser, W.G., E.H. Ross, and R.F. Anderson, (1981) Seasonality in the supply of sediment to the deep Sargasso Sea and implications for the rapid transfer of matter to the deep ocean, Deep-Sea Research, 28A (5), 495-505.

[7] Deuser, W.G., T.D. Jickells, P. King, and R.F. Commeau, (1995) Decadal and annual change in biogenic opal and carbonate fluxes to the deep Sargasso Sea, Deep-Sea Research I, 42 (11/12), 1923-1932.

[8] Dymond, J., and R. Collier, (1988) Biogenic particle fluxes in the equatorial Pacific: Evidence for both high and low productivity during the 1982-1983 El Nino, Global Biogeochemical Cycles, 2 (2), 129-137.

[9] Dymond, J., Lyle, M., (1985) Flux comparisons between sediments and sediment traps in the eastern tropical pacific: Implications for atmospheric CO_2 variations during the Pleistocene. Limnology and Oceanography. 30 (4), 699 - 712.

[10] Dymond, J., and M. Lyle, (1994) Particle fluxes in the ocean and implications for sources and preservation of ocean sediments, in Studies in Geophysics: Material Fluxes on the Surface of the Earth, edited by NRC, pp. 125-142, National Academy Press, Washington, D.C..

[11] Fahl, K., Nothig, E-M., (2007) Lithogenic and biogenic particle fluxes on the Lomonosov Ridge (central Arctic Ocean) and their relevance for sediment accumulation: vertical vs. lateral transport. Deep Sea Research Part I: Oceanographic Research Papers. 54 (8), 1256 - 1272.

[12] Fischer, G., and G. Wefer, (1996) Seasonal and interannual particle fluxes in the eastern equatorial Atlantic from 1989 to 1991: ITCZ migrations and upwelling, in Par-

ticle Flux in the Ocean, edited by V. Ittekkot, P. Schafer, S. Honjo, and P.J. Depetris, pp. 199-214, John Wiley & Sons, New York, N.Y..

[13] Fischer, G., D. Futterer, R. Gersonde, S. Honjo, D. Ostermann, and G. Wefer, (1988) Seasonal variability of particle flux in the Weddell Sea and its relation to ice cover, Nature, 335 (6189), 426-428, 1988.

[14] Fischer, G., V. Ratmeyer, and G. Wefer, (2000) Organic carbon fluxes in the Atlantic and the Southern Ocean: relationship to primary production compiled from satellite radiometer data, Deep-Sea Research II, 47, 1961-1997.

[15] Fischer, G., R. Gersonde, and G. Wefer, (2002) Organic carbon, biogenic silica and diatom fluxes in the marginal winter sea-ice zone and the Polar Front Region: interannual variations and differences in composition, Deep-Sea Research II, 49, 1721-1745.

[16] Fischer, G., Wefer, G., Romero, O.E., Dittert, N., Donner, B., Ratmeyer, V. (2004), Transfer of particles into the deep atlantic and the global ocean: control of nutrient supply and ballast production. In: Wefer, G., Mulitza, S., Ratmeyer, V. (eds), The South Atlantic in the late Quaternary: Reconstructins of Material Budgets and Current Systems. Springer: Berlin, pp. 21-46

[17] Haake, B., V. Ittekkot, T. Rixen, V. Ramaswamy, R.R. Nair, and W.B. Curry, (1993) Seasonality and interannual variability of particle fluxes to the deep Arabian Sea, Deep-Sea Research I, 40 (7), 1323- 1344.

[18] Haake, B., T. Rixen, T. Reemtsma, V. Ramaswamy, and V. Ittekkot, (1996) Processes determining seasonality and interannual variability of settling particle fluxes to the deep Arabian Sea, in Particle Flux in the Ocean, edited by V. Ittekkot, P. Schafer, S. Honjo, and P.J. Depetris, pp. 251-270, John Wiley & Sons, New York, N.Y.

[19] Handa, N., A. Hosokawa, and T. Masuzawa, (1997) Studies on vertical flux of particulate carbon and silica in the ocean and factors controlling the chemical fluxes, in Biogeochemical Processes in the North Pacific, edited by S. Tsunogai, pp. 249-255, Japan Marine Science Foundation, Tokyo.

[20] Hebbeln, D., M. Marchant, and G. Wefer, (2000) Seasonal variations of the particle flux in the Peru-Chile current at 30°S under 'normal' and El Nino conditions, Deep-Sea Research II, 47, 2101-2128.

[21] Heggie, D., C. Maris, A. Hudson, J. Dymond, R. Beach, and J. Cullen, (1987) Organic carbon oxidation and preservation in NW Atlantic continental margin sediments, in Geology and Geochemistry of Abyssal Plains, edited by P.P.E. Weaver, and J. Thomson, pp. 215-236, Geological Society Special Publication.

[22] Honda, M.C., (2001) Studies of carbon cycles in the NW Pacific Ocean by sediment trap and C14 data, PhD thesis, Hokkaido University, Sapporo, Japan, (Part 2 of 2)

[23] Hong, G.H., S.-M. Choe, M.-S. Suk, J.-Y. Na, I.C. Sin, C.S. Chung, and S.H. Kim, (1996) Annual biogenic particle fluxes to the interior of the East Japan Sea, a large marginal sea of the Northwest Pacific, in Proceedings of the International Marine Science Symposium, edited by S. Tsunogai, pp. 300-321, Japan Marine Science

Foundation, Mutsu, Aomori, Japan.

[24] Honjo, S., (1982) Seasonality and interaction of biogenic and lithogenic particulate flux at Panama Basin, Science, 218, 883-884.

[25] Honjo, S., (1990) Particle fluxes and modern sedimentation in the Polar Oceans, in Polar Oceanography, Part B: Chemistry, Biology, and Geology, edited by W.O. Smith, pp. 687-739, Academic Press, New York.

[26] Honjo, S., (1996) Fluxes of particles to the interior of the open oceans, in Particle Flux in the Ocean, edited by V. Ittekkot, P. Schafer, S. Honjo, and P.J. Depetris, pp. 91-154, John Wiley & Sons, New York, N.Y..

[27] Honjo, S., and S.J. Manganini, (1993) Annual biogenic particle fluxes to the interior of the North Atlantic Ocean; studied at 34°N 21°W and 48°N 21°W, Deep-Sea Research I, 40 (1-2), 587-607.

[28] Honjo, S., S. Manganini, A.L. Karowe, and B.L. Woodward, (1987) Particle fluxes, northeastern Nordic Seas: 1983-1986, in WHOI Technical Report, pp. 1-84, WHOI, Woods Hole, MA.

[29] Honjo, S., S. J. Manganini, G. Wefer, (1988) Annual particle flux and a winter outburst of sedimentation in the northern Norwegian Sea., Deep-Sea Research 35: pp.1223-1234.

[30] Honjo, S., D. W. Spencer, W. D. Gardner, (1992) Sediment trap intercomparison experiment in the Panama Basin, 1979, Deep Sea Research. Part A: Oceanographic Research Papers, vol.39, no.2, pp.333-358.

[31] Honjo, S., J. Dymond, R. Collier, and S.J. Manganini, (1995) Export production of particles to the interior of the equatorial Pacific Ocean during the 1992 EqPac experiment, Deep-Sea Research II, 42 (2-3), 831- 870.

[32] Honjo, S., Dymond, J., Prell, W., Ittekkot, V., (1999) Monsoon-controlled export fluxes to the interior of the Arabian Sea. Deep-Sea Research Part II - Topical Studies in Oceanography. 46 (8-9), 1859 - 1902.

[33] Honjo, S., R. Francois, S. Manganini, J. Dymond, and R. Collier, (2000) Particle fluxes to the interior of the Southern Ocean in the western Pacific sector along 170°W, Deep-Sea Research II, 47, 3521-3548.

[34] Ittekkot, V., R.R. Nair, S. Honjo, V. Ramaswamy, M. Bartsch, S. Manganini, and B.N. Desai, (1991) Enhanced particle fluxes in Bay of Bengal induced by injection of fresh water, Nature, 351, 385-387.

[35] Jickells, T.D., P.P. Newton, P. King, R.S. Lampitt, and C. Boutle, (1996) Comparison of sediment trap records of particle fluxes from 19°to 48°N in the northeast Atlantic and their relation to surface water productivity, Deep Sea Research I, 43 (7), 971-986.

[36] Kawahata, H., (2002) Suspended and settling particles in the Pacific, Deep-Sea Research II, 49, 5647-5664.

[37] Kawahata, H., A. Suzuki, and H. Ohta, (1998) Sinking particles between the equatorial and subarctic regions (0°N 46°N) in the central Pacific, Geochemical Journal, 32, 125-133.

[38] Kawahata, H., A. Suzuki, and H. Ohta, (2000) Export fluxes in the western Pacific warm pool, Deep-Sea Research I, 47, 2061-2091.

[39] Kempe, S., and H. Knaack, (1996) Vertical particle flux in the western Pacific below the north equatorial current and the equatorial counter current, in Particle Flux in the Ocean, edited by V. Ittekkot, P. Schafer, S. Honjo, and P.J. Depetris, pp. 313-323, John Wiley & Sons, New York, N.Y.

[40] Lakashin, V.N., (2008) Sedimentary matter fluxes in the contour current sedimentation system over the continental slope of the Norwegian Sea. Oceanology. 48 (5), 701 - 709.

[41] Lyle, M., R. Zahn, F. Prahl, J. Dymond, R. Collier, N. Pisias, and E. Suess, (1992) Paleoproductivity and carbon burial across the California Current: the multitracers transect, 42°N, Paleoceanography, 7 (3), 251-272.

[42] Nair, R.R., V. Ittekkot, S.J. Manganini, V. Ramaswamy, B. Haake, E.T. Degens, B.N. Desai, and S. Honjo, (1989) Increased particle flux to the deep ocean related to monsoons, Nature, 338, 749-751.

[43] Neuer, S., V. Ratmeyer, R. Davenport, G. Fischer, and G. Wefer, (1997) Deep water particle flux in the Canary Island region: seasonal trends in relation to long-term satellite derived pigment data and lateral sources, Deep-Sea Research I, 44 (8), 1451-1466.

[44] Pudsey, C.J., and P. King, (1997) Particle fluxes benthic processes and the paleoenvironmental record in the northern Weddell Sea, Deep-Sea Research I, 44, 1841-1876.

[45] Schafer, P., V. Ittekkot, M. Bartsch, R.R. Nair, and J. Tiemann, (1996) Fresh water influx and particle flux variability in the Bay of Bengal, in Particle Flux in the Ocean, edited by V. Ittekkot, P. Schafer, S. Honjo, and P.J. Depetris, pp. 271-292, John Wiley & Sons, New York, N.Y.

[46] Scholten, J.C., J. Fietzke, S. Vogler, M.M. Rutgers van der Loeff, A. Mangini, W. Koeve, J. Waniek, P. Stoffers, A. Anita, and J. Kuss, (2001) Trapping efficiencies of sediment traps from the deep eastern North Atlantic: the 230Th calibration, Deep-Sea Research II, 48, 2383-2408.

[47] Takahashi, T., N. Fujitani, M. Yanada, and Y. Maita, (1997) Five year long particle fluxes in the central subarctic Pacific and the Bering Sea, in Biogeochemical Processes in the North Pacific, edited by S. Tsunogai, pp. 277-289, Japan Marine Science Foundation, Tokyo.

[48] Takahashi, K., N. Fujitani, M. Yanada, and Y. Maita,(2000) Long-term biogenic particle fluxes in the Bering Sea and the central subarctic Pacific Ocean, 1990-1995, Deep-Sea Research I, 47, 1723-1759.

[49] von Bodungen, B., A. Anita, E. Bauerfeind, O. Haupt, W. Koeve, E. Machado, I. Peeken, R. Peinert, S. Reitmeier, C. Thomsen, M. Voss, M. Wunch, U. Zeller, and

B. Zeitzschel, (1995) Pelagic processes and vertical flux of particles: an overview of a long-term comparative study in the Norwegian Sea and Greenland Sea, Geol. Rundsch, 84, 11-27.

[50] Wefer, G., and G. Fischer, (1991) Annual primary production and export flux in the Southern Ocean from sediment trap data, Marine Chemistry, 35, 597-613.

[51] Wefer, G., and G. Fischer, (1993) Seasonal patterns of vertical particle flux in equatorial and coastal upwelling areas of the eastern Atlantic, Deep-Sea Research I, 40 (8), 1613-1645.

[52] Wong, C.S., F.A. Whitney, D.W. Crawford, K. Iseki, R.J. Matear, W.K. Johnson, J.S. Page, and D. Timothy, (1999) Seasonal and interannual variability in particle fluxes of carbon, nitrogen and silicon from time series of sediment traps at Ocean Station P, 1982-1993: Relationship to changes in subarctic primary productivity, Deep-Sea Research II, 46, 2735-2760.

[53] Treguer unpublished

[54] Honjo unpublished

[54] Francois, R., S. Honjo, R. Krishfield, and S. Maganini (2002) Factors controlling the flux of organic carbon to the bathypelagic zone of the ocean. Global Biogeochemical Cycles. 16 (4)

A.2 Geographically Weighted Regression Model Output

The geographically weighted regression analysis in Chapter 3 was performed using the GWR 3.0 software provided by Martin Charlton at the National Centre for Geocomputation at the National University of Ireland Maynooth. A PDF manual for the software can be found online at:

http://eprints.ncrm.ac.uk/90/1/MethodsReviewPaperNCRM-006.pdf

Since the writing of this thesis, an updated version of the software (GWR 4.0) has been made available. The analysis performed using GWR 3.0 has been checked using GWR 4.0 and is the same. GWR 4.0 is available freely online at:

https://geodacenter.asu.edu/gwr_software

The output from the analysis shown in Figure 3.7 in Chapter 3 using GWR 3.0 is provided here. The output has been abbreviated by removing the casewise diagnostics to focus on the statistical ouput.

```
    **************************************************
    *       Geographically Weighted Regression       *
    *               Release 3.4.3                     *
    *               (31-iii-2007)                     *
    *                                                 *
    *       Martin Charlton, Chris Brunsdon           *
    *            A Stewart Fotheringham               *
    *  (c) National University of Ireland Maynooth    *
    **************************************************
Program starts at: Mon Jul 02 09:08:07 2012

** Program limits:
** Maximum number of variables.....      52
** Maximum number of observations..   80000
** Maximum number of fit locations.   80000

156_model
** Observed data file:          D:\GeographicallyWeightedRegression\SedT
** Prediction location file:  Estimation at sample point locations
** Result output file:          D:\GeographicallyWeightedRegression\SedT

** Variables in the data file...
 long     lat      poc      caco3     opal      litho

** Dependent (y) variable..........poc
** Easting  (x-coord) variable.....long
** Northing (y-coord) variable.....lat
** No weight variable specified
** Independent variables in your model...
 caco3    opal      litho
** Kernel type:  Adaptive
** Kernel shape: Bi-Square
** Bandwidth selection by AICc minimisation
** Use all regression points
** Calibration history requested
** Prediction report requested
** Output estimates to be written to .txt file
** Monte Carlo significance tests for spatial variation
** Casewise diagnostics to be printed

*** Analysis method ***
*** Geographically weighted multiple regression
** Spherical coordinates: Great Circle Distance
****************************************************************
*                                                             *
*       GEOGRAPHICALLY WEIGHTED GAUSSIAN REGRESSION           *
*                                                             *
```

```
*************************************************************
 Number of data cases read: 156
 Observation points read...

 Dependent mean=  1.653247
 Number of observations, nobs= 156
 Number of predictors,   nvar= 3
 Observation Easting extent:   6.21337223
 Observation Northing extent:  2.51711392
 ** NB these are in radians
*Finding bandwidth...
   ... using all regression points
 This can take some time...
 *Calibration will be based on  156 cases
*Adaptive kernel sample size limits:      10     156
 *AICc minimisation begins...
            Bandwidth                      AICc
         55.116481270000         297.183880324985
         83.000000000000         302.634666579220
         37.883518934058         321.201178820903
         65.767037664058         283.443482508653
         72.349443527998         291.796148826405
         61.698887133940         287.104618757539
         68.281292968108         286.947593281877
         64.213142437990         286.395223312413
         66.727397730668         284.829063596995
** Convergence after      9 function calls
** Convergence: Local Sample Size=      66


*************************************************************
*          GLOBAL REGRESSION PARAMETERS                     *
*************************************************************
Diagnostic information...
Residual sum of squares.........        73.524015
Effective number of parameters..         4.000000
Sigma...........................         0.695493
Akaike Information Criterion....       335.758771
Coefficient of Determination....         0.657990
Adjusted r-square...............         0.648930


Parameter              Estimate              Std Err                    T
---------            ------------          ------------            ------------

Intercept        0.012317497019        0.114000329279        0.108047910035
 caco3           0.089002080293        0.006689772651       13.304201126099
 opal            0.022807854623        0.005475704736        4.165282249451
 litho           0.026949395939        0.004912231407        5.486182212830
```

```
************************************************************
*              GWR ESTIMATION                             *
************************************************************
 Fitting Geographically Weighted Regression Model...
 Number of observations............ 156
 Number of independent variables... 4
   (Intercept is variable 1)
 Number of nearest neighbours...... 66
 Number of locations to fit model.. 156

 Diagnostic information...
 Residual sum of squares.........           37.633261
 Effective number of parameters..           20.016103
 Sigma...........................            0.526068
 Akaike Information Criterion....           269.820752
 Coefficient of Determination....            0.824942
 Adjusted r-square...............            0.798984


************************************************************
*                    ANOVA                                *
************************************************************
        Source                   SS        DF        MS           F
OLS Residuals                   73.5      4.00
GWR Improvement                 35.9     16.02     2.2409
GWR Residuals                   37.6    135.98     0.2767      8.0973



************************************************************
*       PARAMETER 5-NUMBER SUMMARIES                      *
************************************************************
   Label      Minimum  Lwr Quartile     Median  Upr Quartile     Maximum
--------  -------------  -------------  -------------  -------------  ----

Intrcept   -0.210956     -0.147014      0.079721      0.271950      0.502543
caco3       0.022056      0.044169      0.066327      0.101922      0.133332
opal       -0.001033      0.023191      0.033392      0.045663      0.156923
litho      -0.065595      0.018597      0.022193      0.050515      0.073473


         <----------------- LOWER ----------------->
   Label Far Out   Outer Fence Outside   Inner Fence
--------  -------  -------------  -------  ------------

Intrcept     0      -1.403906       0      -0.775460
caco3        0      -0.129091       0      -0.042461
opal         0      -0.044222       0      -0.010515
litho        0      -0.077155      19      -0.029279
```

```
            <----------------- UPPER ----------------->
     Label    Inner Fence Outside   Outer Fence Far Out
    --------  ------- -------------  ------- ------

    Intrcept    0.900396      0       1.528842      0
    caco3       0.188552      0       0.275182      0
    opal        0.079369      7       0.113076     25
    litho       0.098391      0       0.146267      0


    **************************************************
    *                                                *
    *   Test for spatial variability of parameters   *
    *                                                *
    **************************************************


    Tests based on the Monte Carlo significance test
    procedure due to Hope [1968,JRSB,30(3),582-598]

    Parameter                   P-value
    ----------            ------------------

    Intercept               0.30000   n/s
    caco3                   0.00000   ***
    opal                    0.00000   ***
    litho                   0.00000   ***


    *** = significant at .1% level
    **  = significant at 1% level
    *   = significant at 5% level


    Program terminates normally at: Mon Jul 02 09:08:08 2012
```

APPENDIX B

Code

This appendix contains code used in Chapter 4 to diagnose a transport matrix in the Earth System Model of Intermediate Complexity: GENIE, and code to process the output in MATLAB. The appendix also contains the MATLAB code for the steady-state model described in Chapter 5.

B.1 Transport Matrix

B.1.1 Context of Transport Matrix in GENIE

The Transport Matrix as described in *Khatiwala et al.* (2005) and *Khatiwala* (2007) contains the finite difference tendency calculated in an ocean model:

$$\frac{d\mathbf{c}}{dt} = \frac{\mathbf{c}^{n+1} - \mathbf{c}^n}{\Delta t} = \mathbf{A}'^n \mathbf{c}^n + \mathbf{q}'^n \tag{B.1}$$

where $\mathbf{A}$' is the transport matrix containing the finite difference tendencies (dt^{-1}), $\mathbf{c}$ is a tracer $(\mathrm{mol\,kg}^{-1})$ and $\mathbf{q}$' is a vector of source/sinks $(\mathrm{mol\,kg}^{-1}\,\mathrm{dt}^{-1})$. The superscripts refer to the time step index. Rearranging Equation B.1 for $\mathbf{c}^{n+1}$:

$$\mathbf{c}^{n+1} = (\mathbf{I} + \mathbf{A}'\Delta t)\mathbf{c}^n + \mathbf{q}'^n \Delta t \tag{B.2}$$

In GENIE the tracer experiments diagnose the tracer distribution resulting from the ocean circulation acting on a unit flux over one time step. Therefore, the GENIE transport matrix $(\mathbf{A})$ described and used in this thesis is equivalent to $(\mathbf{I} + \mathbf{A}'\Delta t)$.

B.1.2 Code

The following code fragments perform and record the colour tracer experiments in GENIE required to diagnose a transport matrix (*Khatiwala et al.*, 2005). The code fragments are described below, are written in Fortran 90 and run within the BIOGEM module:

- *biogem.f90*: The code occurs in two parts. The second part (MUFFIN MATRIX PART II) initialises a single ocean grid-box per iteration of the code with 1 mol kg^{-1} of the 'blue' colour tracer. This is then passed by BIOGEM to the GOLD-STEIN module. In the following call of the BIOGEM routine during the next time-step, the first part of the code (MUFFIN MATRIX PART I) loops over the 'blue' colour tracer array, having returned from GOLDSTEIN, recording any non-zero values. The value of the tracer in the array is written to a text file along with grid-box indices. The code takes advantage of the vectorisation of the tracer array already in cGENIE.

- *biogem_lib.f90*: Variables used for indexing and looping are declared here.

- *initialise_biogem.f90*: The index and loop variables are set to initial values here.

To diagnose a transport matrix:

1. A spin-up run should first be run with the 'blue' colour tracer selected in the base-config file. The spin-up should be run as long as required or when a desired steady state ocean circulation has been reached. A recommended run time is 10,000 years. The spin-up is then used as a restart file for the run where the transport matrix will be diagnosed.

2. Code should then be copied into the *biogem.f90*, *biogem_lib.f90* and *initialise_biogem.f90*.

3. Start a new run, using the spin-up as a restart, using the same base and user configuration files as the restart. The time-stepping ratio between BIOGEM and GOLDSTEIN should be set to 1 to represent the time-step taken by GOLDSTEIN.

4. An output file will appear in the experiment output file which can be processed by MATLAB using the *sparse* function.

All code and line number references are correct for the SVN revision 9167 of the muffin branch of cGENIE. Please note this code is not available via the SVN repository.

B.1.2.1 biogem.f90 code

The code sits within the biogem_tracercoupling subroutine. The variables are first declared:

```
!muffin matrix variables
integer::loop_count
real::loc_colb,loc_V
character(len=127)::loc_filename
```

The following code loops over the colour tracer array from the previous time-step, recording the concentration and grid-box indices if the concentration is > 0. This is then written to a text file and the corresponding grid-boxes set to a concentration of 0:

```
!!! ************************ !!!
!!! MUFFIN MATRIX PART I !!!
!!! ************************ !!!
! find grid-boxes with colour tracer in from ts (from goldstein)

! need to stagger this after the initialisation step (PART II)
if(matrix_loop.gt.1)THEN
print*,'<<<<Recovering Matrix Information'
! set-up unique filename for writing data n.b.
! matrix_loop will always be matrix_loop-1 to correspond to prev. initialisation
loc_filename='muffin_matrix'
open(22,file=loc_filename,position='append')
! main loop over colb in ts
loop_count=1
do n=1,n_vocn,1
loc_k1 = loc_vts(n)%k1
do k=n_k,loc_k1,-1
! see if anything is sitting in grid-boxes (as mol kg-1) and if so write data
loc_colb=loc_vts(n)%mk(io2l(io_colb),k)!*vphys_ocn(n)%mk(ipo_M,k)

if(abs(loc_colb).gt.const_real_nullsmall)THEN
PRINT*,'found stuff in:',matrix_loop-1,loc_vts(n)%i,loc_vts(n)%j,k,loc_colb

write(22,FMT='(I4,4x,I4,4x,e13.7)') &
    &matrix_loop-1,&
    &loop_count,&
    &loc_colb

end if
loop_count=loop_count+1
end do

end do
close(22) ! close file
END if

! get out clause when all boxes are initialised....
! n_vocn is not the number of all wet
! grid boxes though-> is the number of wet 2d grid points!
IF(matrix_vocn_n.gt.n_vocn)THEN

! write out an indexing file
open(21,file='muffin_matrix_v_index')
```

```
do n=1,n_vocn,1
loc_k1 = loc_vts(n)%k1
do k=n_k,loc_k1,-1

write(21,FMT='(A2,1X,A2,1X,A2)')&
&fun_conv_num_char_n(2,loc_vts(n)%i),   &
&fun_conv_num_char_n(2,loc_vts(n)%j),   &
&fun_conv_num_char_n(2,k)

end do
end do
close(21)
! shut down the simulation
print*, '*** <<< ALL MATRIX BOXES CALCULATED.....STOP!'
stop
end if



!!! ******************************** !!!
!!! END OF MUFFIN MATRIX PART I !!!
!!! ******************************** !!!
```

After recording the results of the last time-step, this code initialises the next grid-box with a unit of colour tracer:

```
!!! ************************* !!!
!!! MUFFIN MATRIX PART II !!!
!!! ************************* !!!
! initialise colour tracer in ts (going to goldstein)
! n.b. matrix_count & matrix_k are set in biogem_lib
!IF(par_misc_matrix)THEN
loc_k1=loc_vts(matrix_vocn_n)%k1
PRINT*,'<<<< Initialising matrix - Box Number:',&
& matrix_loop,'Box i:',loc_vts(matrix_vocn_n)%i,&
& 'Box j:',loc_vts(matrix_vocn_n)%j,&
& 'Box k:',matrix_k,'loc k1',loc_k1

! initialising grid_box with 1 mol kg-1 of colour tracer
loc_vts(matrix_vocn_n)%mk(io2l(io_colb),matrix_k)=&
& 1!*vphys_ocn(matrix_vocn_n)%mk(ipo_rM,matrix_k)

IF(matrix_k.gt.loc_k1)THEN ! if matrix_k > the bottom....
matrix_k=matrix_k-1 ! decrement matrix_k
matrix_loop=matrix_loop+1 ! increment matrix_loop
else ! then matrix_k is equal to the bottom....
matrix_k=n_k ! reset matrix_k
matrix_loop=matrix_loop+1 ! increment matrix_loop
```

```
matrix_vocn_n=matrix_vocn_n+1 ! increment matrix_vocn_count
end if

!
!!! ********************************* !!!
!!! END OF MUFFIN MATRIX PART II !!!
!!! ********************************* !!!
```

```
integer::matrix_vocn_n
integer::matrix_k
integer::matrix_loop
NAMELIST /ini_biogem_nml/matrix_vocn_n,matrix_k,matrix_loop
```

B.1.2.3   initialise_biogem.f90 code

```
matrix_vocn_n=1
matrix_k=8
matrix_loop=1
```

B.2   Steady-state Model Code

B.2.1   MATLAB Code

B.2.1.1   OMFG.m

```matlab
function [ Model_Output ] = OMFG (...
    A,...
    runtime,...
    dt_ratio,...
    l_POC1,...
    frac2,...
    uptake_scheme,...
    restart,...
    model_directory,...
    save_output)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Offline Matrix-based Fast GENIE (OMFG)
% "oh my f*cking god that's fast"
%
% forward simulation of phosphate only matrix-genie
%
% govering equations:   dPO4/dt=A*PO4-Jup+Jremin+(k*DOP)
%                       dDOP/dt=A*DOP+(DOP_frac*Jup)-(k*DOP)
```

```matlab
%
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% initialise variables/arrays

% -- load in genie-based data if needed
addpath(['C:\Users\ggzjw\Local Documents\...
    Matlab_Files\Spatial_Remin\' model_directory])
load('matrix_vars')
load('genie_vars')
if ¬strcmp(uptake_scheme,'MM')
    load(['C:\Users\ggzjw\Local Documents\Matlab_Files\Spatial_Remin\ODV_PO4']);
end

% -- basic
dt=1/96*dt_ratio;  % genie timestep (years)
conv_d_s=24*3600;
conv_d_yr=365.25;
conv_yr_s=365.25*24.0*3600.0;
z0=zt_edges(2); % bottom of surface layer (m)
loop_count=1;
dt_loop_count=1;
solconst=1368;  % solar constant (W m-2)

% -- biology relevant parameters
l_POC2=1000000.0;  % efolding depth of refractory POP (m)
DOP_frac=0.66;    % fraction of export for DOP
rDOP_frac=1-DOP_frac;   % reciprical of DOP_frac
DOP_k=(1/0.5);  % DOP decay rate (yr-1 -> year-1)
tau=30/conv_d_yr;  % restoring timescale (days -> yrars)
r_tau=1/tau;   % reciprical restoring timescale (1/years)
u0PO4=1.9582242E-06;   % maximum PO4 uptake (mol kg-1 yr-1)
KPO4=2.1989611E-07;   % Michaelis-Menton half sat. conc. (mol kg-1);

% -- arrays
J=zeros(nb,2);  % source/sink array
PO4=zeros(nb,2);    % tracer array at current time step, (:,1)=PO4 (:,2)=DOP
PO4_1=zeros(nb,2);   % tracer array at previous time step
JDOP=zeros(nb,1);     % DOP remineralisation array (mol kg-1)
dPO4_dt=zeros(nb,1); % dPO4/dt
remin_fracs=zeros(8,36,36);
Ib=find(v_index.k==8);
prod_mod=zeros(nb,1);
prod_mod=(1-(SEAICE./100)).*(SOLFOR./solconst)*u0PO4;
Juptake=zeros(nb,1);
Jremin=zeros(nb,1);


Model_Output.Year=zeros(1,numel(save_output));
Model_Output.PO4=zeros(nb,numel(save_output));
Model_Output.DOP=zeros(nb,numel(save_output));
Model_Output.PO4_Uptake=zeros(nb,numel(save_output));
Model_Output.PO4_Remin=zeros(nb,numel(save_output));
Model_Output.DOP_Remin=zeros(nb,numel(save_output));
```

```matlab
Model_Output.PO4_Inventory=zeros(1,numel(save_output));
Model_Output.DOP_Inventory=zeros(1,numel(save_output));

%% remin matrix
[remin]=create_remin_matrix(l_POC1,l_POC2,frac2);

%% run forward simulation

% initialise tracer arrays
if restart==1
    [matrix_PO4,matrix_DOP]=read_genie_netcdf(...
        'fields_biogem_3d.nc',1,v_index,'ocn_PO4','ocn_DOM_P');
    PO4_1(:,1)=matrix_PO4;
    PO4_1(:,2)=matrix_DOP;
else
    PO4_1(:,1)=2.159/1e6;
% initial PO4 (umol kg-1 -> mol kg-1) (Ridgwell et al. 2007)
    PO4_1(:,2)=0.0235/1e6;
% inital DOP umol kg-1 -> mol kg-1) (Najjar et al. 2007)
end



for t=1:96*runtime

    % ocean + biology step
    if dt_loop_count==dt_ratio || dt_ratio==1

    % reset source/sink arrays
    J=zeros(nb,2); Juptake=zeros(nb,1); Jremin=zeros(nb,1);

    % DOP remineralisation (mol kg-1),
    % n.b. calc seperately to minimise number of calculations
    JDOP=PO4_1(:,2)*DOP_k*dt;

    % Uptake mol kg-1
    if strcmp(uptake_scheme,'MM')
    Juptake(Ib,1)=(dt*(prod_mod(Ib,1).*(PO4_1(Ib,1)./(PO4_1(Ib,1)+KPO4))));
    else
    Juptake(Ib,1)=(dt*(r_tau.*(PO4_1(Ib,1)-ODV_PO4(Ib,1))));
    Juptake(Juptake>0)=0;
    end

    % DOP from uptake, mol kg-1
    J(Ib,2)=(Juptake(Ib,1)*DOP_frac);

    % remineralisation of uptake, mol kg-1
    Jremin(:,1)=(remin*(Juptake*rDOP_frac));

    % DOP remin to PO4, mol kg-1
    J(:,1)=J(:,1)+Jremin-Juptake+JDOP;
    % DOP remin from DOP, mol kg-1
    J(:,2)=J(:,2)-JDOP;

    % transport divergence plus source/sink
    PO4=A*PO4_1+J;
```

```matlab
        dt_loop_count=1;

    % ocean only step
    else
    PO4=A*PO4_1;
    dt_loop_count=dt_loop_count+1;
    end

    % set previous time-step array PO4_1
    PO4_1=PO4;

    % integrate save output
    if (t≥save_output*96-48 & t≤save_output*96+48) % is between save dt=1
        save_ind=(t≥save_output*96-48 & t≤save_output*96+48);

        Model_Output.PO4(:,save_ind)=Model_Output.PO4(:,save_ind)+PO4(:,1);
        Model_Output.DOP(:,save_ind)=Model_Output.DOP(:,save_ind)+PO4(:,2);
        Model_Output.PO4_Uptake(:,save_ind)=...
            Model_Output.PO4_Uptake(:,save_ind)+Juptake;
        Model_Output.PO4_Remin(:,save_ind)=...
            Model_Output.PO4_Remin(:,save_ind)+Jremin;
        Model_Output.DOP_Remin(:,save_ind)=...
            Model_Output.DOP_Remin(:,save_ind)+JDOP;
        Model_Output.PO4_Inventory(:,save_ind)=...
            Model_Output.PO4_Inventory(:,save_ind)+sum(M.*PO4(:,1));
        Model_Output.DOP_Inventory(:,save_ind)=...
            Model_Output.DOP_Inventory(:,save_ind)+sum(M.*PO4(:,2));

    end



end

% output variables
Model_Output.Year=save_output;
Model_Output.PO4=Model_Output.PO4./96;
Model_Output.DOP=Model_Output.DOP./96;
Model_Output.PO4_Inventory=Model_Output.PO4_Inventory./96;
Model_Output.DOP_Inventory=Model_Output.DOP_Inventory./96;

end

%% SUB FUNCTIONS
function [ remin_matrix ] = create_remin_matrix ( l_POC1 , l_POC2 , frac2 )

load('matrix_vars')
load('genie_vars','zt_edges','ocn_mask','M_ratio')

z0=zt_edges(2);

%% set up remin profiles
% initialise an idealised water column, extend to full genie grid
% then adjust for bottom depths satisfying closed system
```

```matlab
if numel(l_POC1)==1
    l_POC1=repmat(l_POC1,[numel(Ib),1]);
end

if numel(frac2)==1
    frac2=repmat(frac2,[numel(Ib),1]);
end

for n=1:numel(Ib)
remin_profile=1-...
    ((1-frac2(n,1))*(exp((z0-zt_edges(2:end))/l_POC1(n,1)))...
    +frac2(n,1)*(exp((z0-zt_edges(2:end))/l_POC2)));
% cumulative remineralisation fractions (for both fractions)

remin_profile_diff=diff([remin_profile(1:end-1) ; 1]);
% remin proportions in each grid-box
remin_profile_diff=[0 ; remin_profile_diff];
% remin proportions in each grid-box

remin_fracs(:,v_index.j(Ib(n)),v_index.i(Ib(n)))=remin_profile_diff;
% extend to full 8x36x36 grid
end
% loop over water columns and adjust remin profiles for bottom depths
for ii=1:36
    for jj=1:36

        loc_mask=ocn_mask(:,ii,jj);
        if nansum(loc_mask)<8
% if not full 8 box depth profile
            loc_profile=remin_fracs(:,ii,jj);
            loc_zbot=find(loc_mask==1,1,'last');
% find wet grid-box above sediment
            remin_fracs(loc_zbot,ii,jj)=loc_profile(loc_zbot)+...
                sum(loc_profile(loc_zbot+1:end));  % add rest of remin to grid-box
        end

    end
end
remin_fracs=f2v(remin_fracs,v_index.i,v_index.j,v_index.rk);
% vectorise
remin_fracs=remin_fracs.*M_ratio;
% combine with ratio between surface mass and grid-box mass

clear remin_profile remin_profile_diff loc_profile loc_zbot loc_mask
% only needed once so clear from memory


%% create remin transformation matrix
% to transform Jup to vector for multiplication with Jremin

water_col_remin=sparse(6210,6210);

nbb=numel(find(v_index.k==8));
Ib1=[find(v_index.k==8);6211];
```

```matlab
for n=2:nbb+1

    temp=zeros(nb,1);
    temp(Ib1(n-1):Ib1(n)-1)=remin_fracs(Ib1(n-1):Ib1(n)-1);

    water_col_remin(:,Ib1(n-1):Ib1(n)-1)=repmat(temp,[1 numel(Ib1(n-1):Ib1(n)-1)]);

end

remin_matrix=water_col_remin;

clear Ib1


end
```

APPENDIX C

Supplementary Figures and Tables

This appendix contains supplementary Tables and Figures for all chapters.

Table C.1: Minimum and maximum overturning circulations for each ensemble member from *Annan et al.* (2005) and for the standard GENIE configuration corresponding to the SYN dataset (*Ridgwell and Hargreaves*, 2007) calculated at the end of the 10000 year spin-up.

Ens.	Global Min.	Global Max.	Atl. Min.	Atl. Max
1	-44.984	44.424	-1.145	20.013
2	-42.595	41.983	-1.009	20.767
3	-48.919	48.096	-0.471	17.1
4	-37.221	36.743	-0.781	17.913
5	-40.909	37.746	-0.965	18.436
6	-44.138	41.438	-0.964	18.519
7	-41.648	39.93	-1.199	20.104
8	-33.106	30.128	-0.773	16.518
9	-37.825	35.532	-0.968	16.697
10	-38.763	38.677	-0.976	17.562
11	-45.308	43.577	-0.918	17.668
12	-43.51	41.387	-1.12	17.875
13	-36.348	32.935	-0.91	16.756
14	-40.963	38.4	-0.918	17.543
15	-35.481	33.514	-0.771	16.994
16	-36.058	35.122	-1.213	19.245
17	-45.322	40.9	-1.287	18.462
18	-36.093	33.451	-1.123	17.622

Ens.	Global Min.	Global Max.	Atl. Min.	Atl. Max
19	-43.719	41.425	-0.97	17.7
20	-39.29	36.775	-1.02	18.446
21	-46.585	43.208	-0.946	17.063
22	-46.512	40.972	-1.605	22.607
23	-44.881	41.41	-0.891	16.35
24	-43.417	42.392	-0.888	15.656
25	-41.159	37.946	-0.979	17.759
26	-38.099	36.185	-1.279	18.826
27	-40.831	37.522	-0.907	16.188
28	-36.051	32.139	-1.267	17.629
29	-41.826	39.273	-1.015	18.428
30	-39.999	37.343	-0.883	16.047
31	-44.028	43.078	-1.128	18.89
32	-40.015	38.424	-0.946	17.374
33	-45.535	41.939	-0.975	17.5
34	-37.05	33.947	-0.933	18.752
35	-46.752	43.682	-2.142	19.603
36	-40.286	37.239	-2.18	18.173
37	-44.404	42.344	-0.639	16.259
38	-38.082	37.064	-1.052	16.523
39	-39.32	36.908	-2.129	17.893
40	-34.405	32.641	-1.369	18.476
41	-38.175	38.101	-0.912	19.18
42	-39.147	37.322	-1.099	18.696
43	-45.06	40.854	-1.08	19.603
44	-43.422	42.015	-1.049	18.123
45	-50.089	44.883	-0.997	18.595
46	-40.868	36.177	-0.942	19.541
47	-42.207	39.9	-0.51	15.787
48	-36.276	32.168	-2.061	18.242
49	-41.262	37.757	-0.995	17.64
50	-42.064	39.877	-0.827	16.787
51	-41.28	37.355	-1.025	16.997
52	-39.448	37.092	-1.142	17.869
53	-36.709	32.89	-0.705	15.349
54	-42.012	39.44	-1.308	17.956
SYN	-44.751	40.226	-0.177	18.221

Table C.2: Parameters values from the GENIE ensemble *Annan et al.* (2005) used in this study. The parameter values for the standard tuned model set-up in *Ridgwell and Hargreaves* (2007) are also shown.

Parameter	Reference	Min.	Max.	SYN
Ocean				
temp0	go_10	10.115	24.904	0.000
temp1	go_11	10.115	24.904	0.000
rel	go_12	0.9	0.9	0.9
sc_tau	go_13	1.078	2.320	1.932
iso	go_14	2480.917	6496.541	4488.812
dia	go_15	5.13×10^{-6}	5.14×10^{-5}	2.72×10^{-5}
inverse minimum drag	go_16	1.853	4.565	2.940
Atmosphere				
atm. Diff. amp. T	ea_12	3045602	5357693	4667550
atm. Diff. amp. Q	ea_13	1274689	2162105	1097107
dist 'n' width	ea_14	0.923	2.005	1.083
slope	ea_15	4.89×10^{-3}	1.48×10^{-1}	6.33×10^{-2}
T_z	ea_16	-2.40×10^{-3}	2.30×10^{-1}	1.12×10^{-1}
q_z	ea_18	4.21×10^{-2}	2.30×10^{-1}	2.27×10^{-1}
q_m	ea_19	4.21×10^{-2}	2.30×10^{-1}	2.27×10^{-1}
extra1a	ea_25	-4.29×10^{-2}	-2.60×10^{-2}	2.12×10^{-2}
extra1b	ea_26	1.48×10^{-1}	2.43×10^{-1}	1.20×10^{-1}
extra1c	ea_27	1.56×10^{-1}	2.58×10^{-1}	1.27×10^{-1}
Ice				
sea-ice eddy diffusivity	gs_11	2023.114	8000	6200
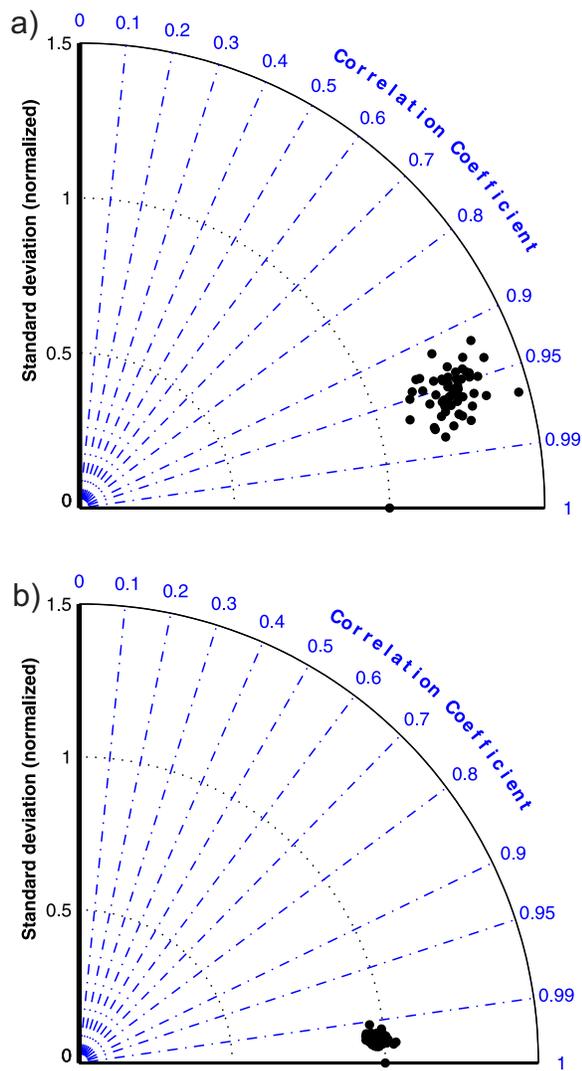
Figure C.1: Taylor diagram comparison of the perturbed physics ensemble (*Annan et al.*, 2005) versus the standard configuration of GENIE (*Ridgwell et al.*, 2007a; *Ridgwell and Hargreaves*, 2007) for (a) salinity and (b) temperature.
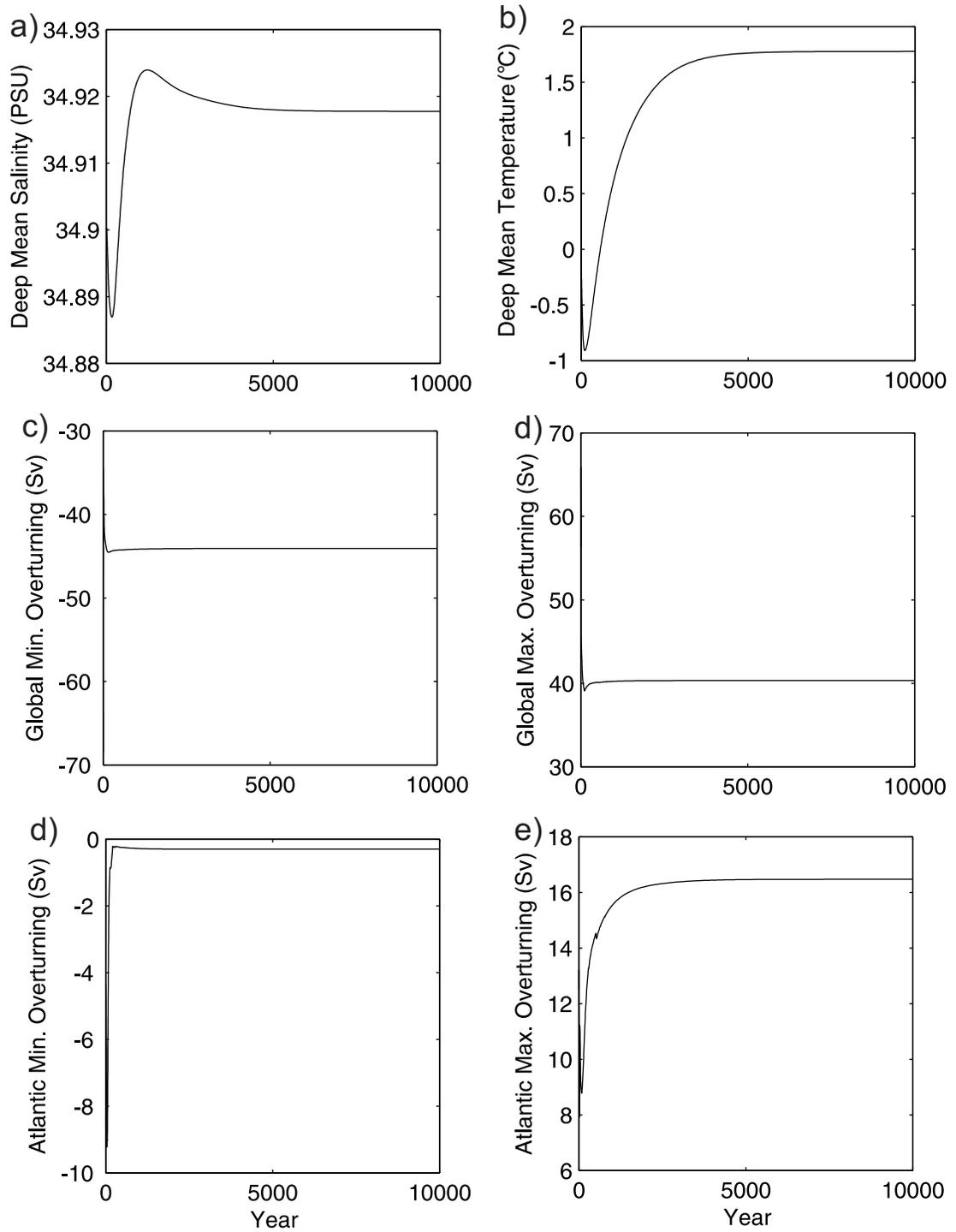
Figure C.2: Time-series of circulation relevant properties during a 10000 year spin up to test for steady state. The spin up is from the run that the SYN dataset is created from (Table 4.2) for (a) deep mean salinity (PSU), (b) deep mean temperature (°C), (c) the global minimum overturning circulation (Sv), (d) the global maximum overturning circulation (Sv), (e) the Atlantic minimum overturning circulation (Sv) and (f) the Atlantic maximum overturning circulation (Sv).
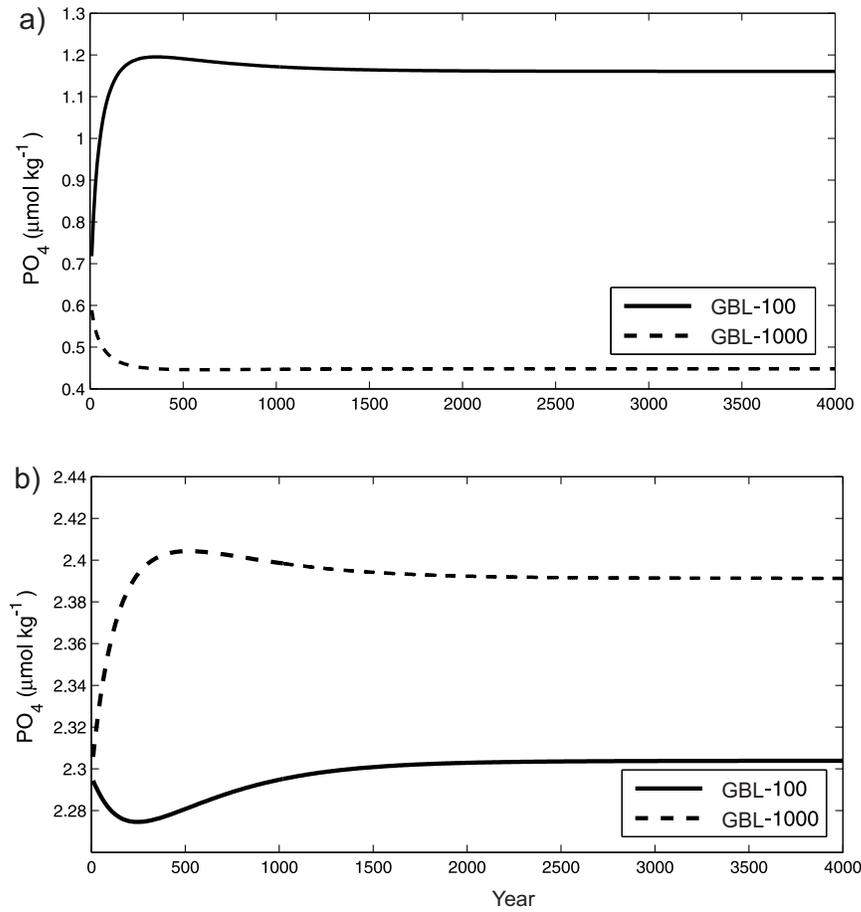
Figure C.3: Transient response of the mean benthic and shallow [PO$_4$] when the global remineralisation depth is perturbed to 100m or 1000m (experiments GBL-100 and GBL-1000). Remineralisation depths are changed at the beginning of the run. Both benthic and surface [PO$_4$] are at equilibrium at 3000 years.
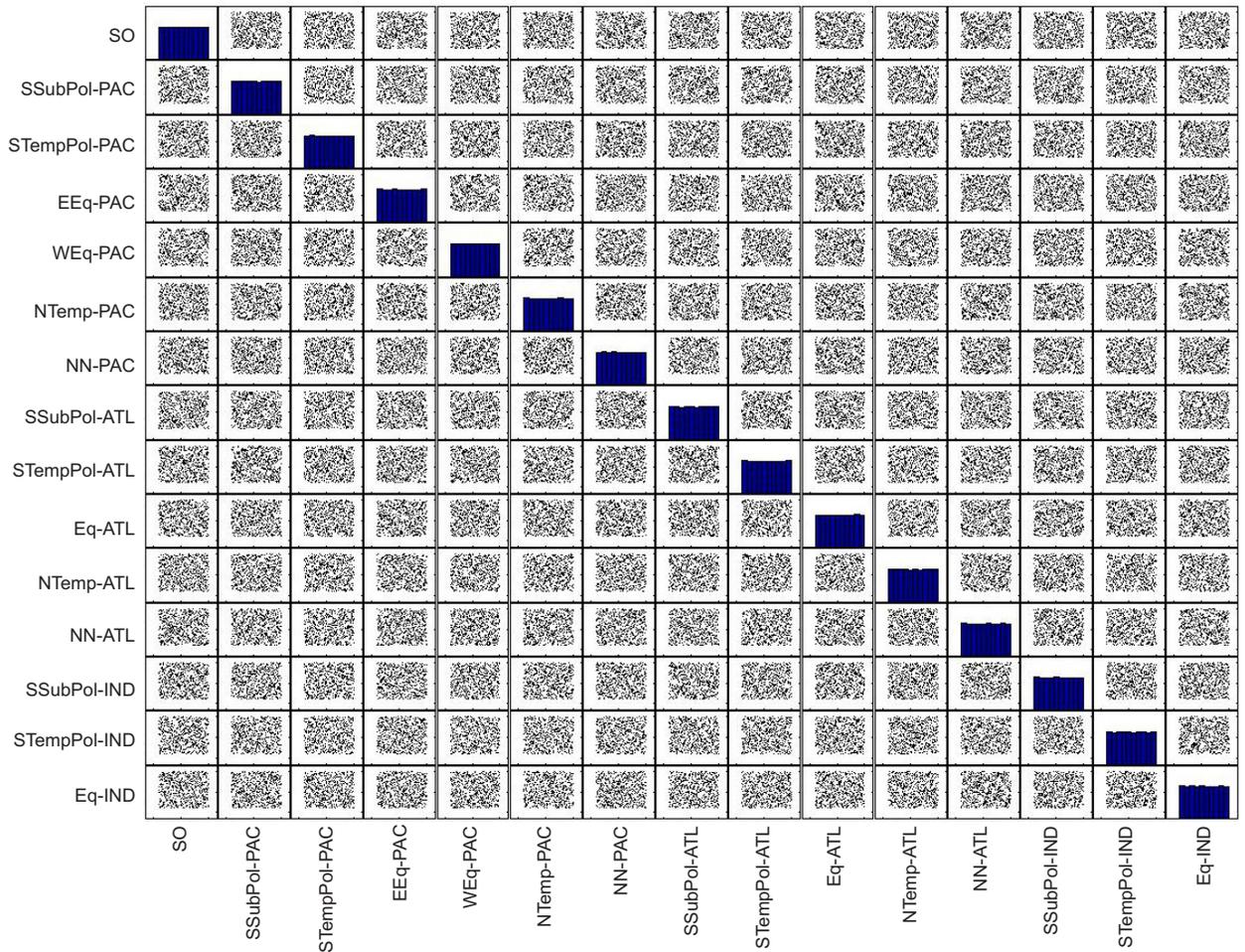
Figure C.4: The sampling of remineralisation depths using a latin hypercube. The 15 regions with 600 samples from a uniform prior distribution between 100m and 1000m are shown. The histograms illustrate the sampling across the prior distribution. Scatter plots of all regions versus each other are also shown. Each scatter plot contains 600 data points.
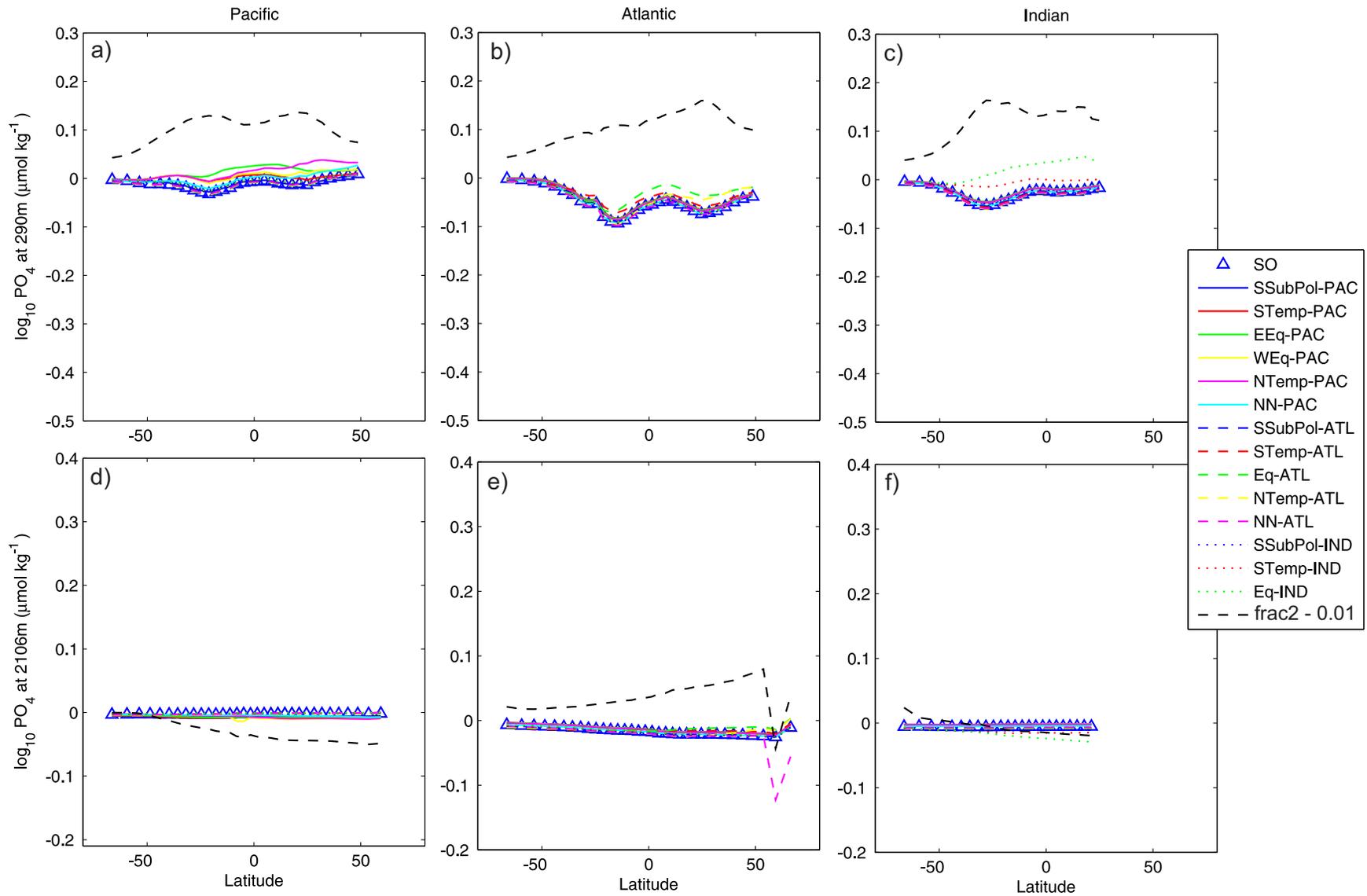
Figure C.5: Sensitivity of [PO$_4$] zonal averages when f^{POP} (Equation 5.5) is decreased. Layout and formatting are identical to Figure 5.6, except that now that f^{POP} is decreased to 0.100 from 0.065, and the remineralisation depth is maintained at 550m
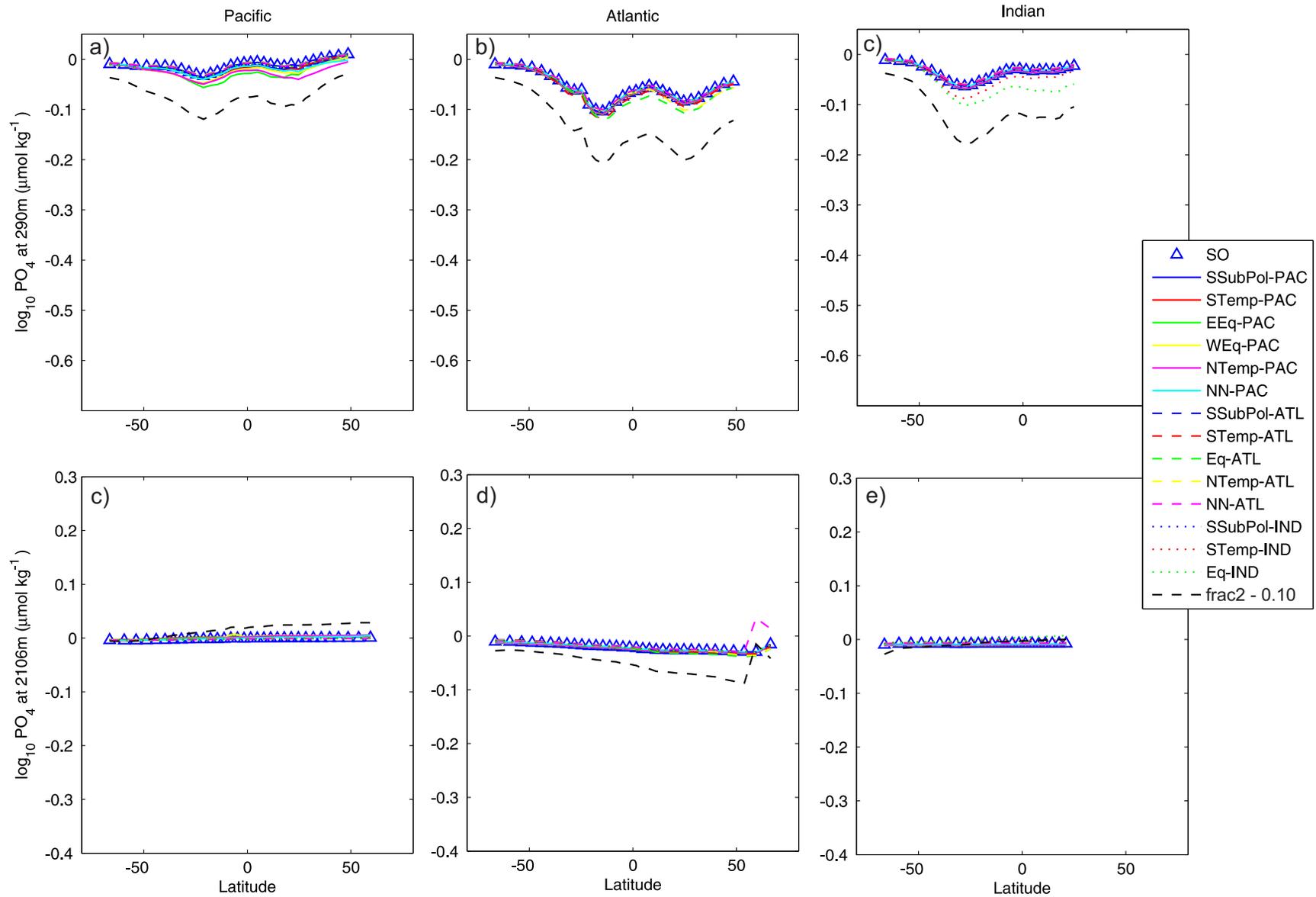
Figure C.6: Sensitivity of [PO$_4$] zonal averages when f^{POP} (Equation 5.5) is increased. Layout and formatting are identical to Figure 5.6, except that now that f^{POP} is increased to 0.100 from 0.065, and the remineralisation depth is maintained at 550m
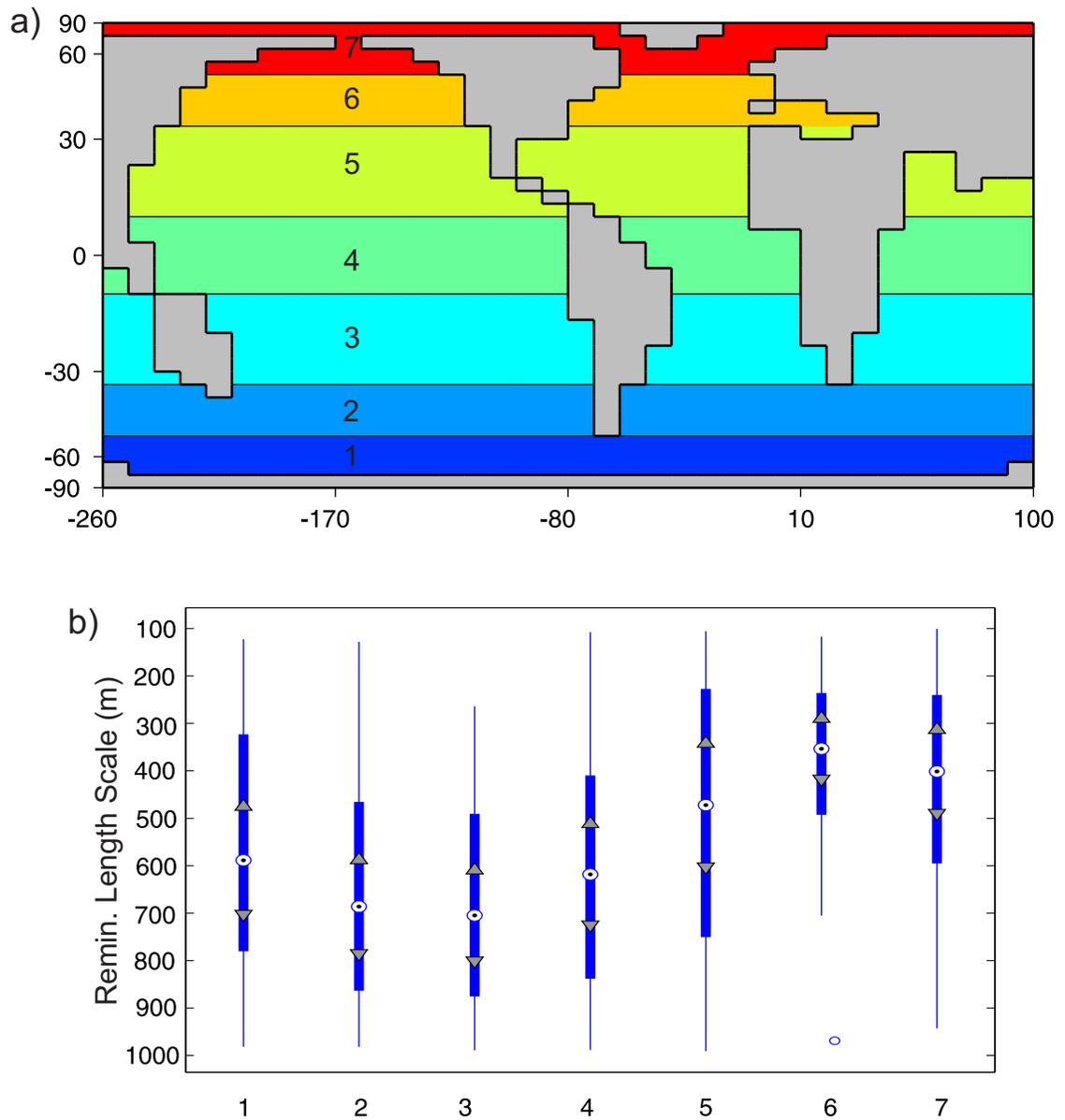
Figure C.7: Regions and optimal solution for n=7 regions. (a) Locations of regions used (b) boxplot of the 40 best solutions for each region comparable to Figure 5.10. 140 Latin hypercube samples were taken, of which 80 had a RMSD smaller than the CTRL run.
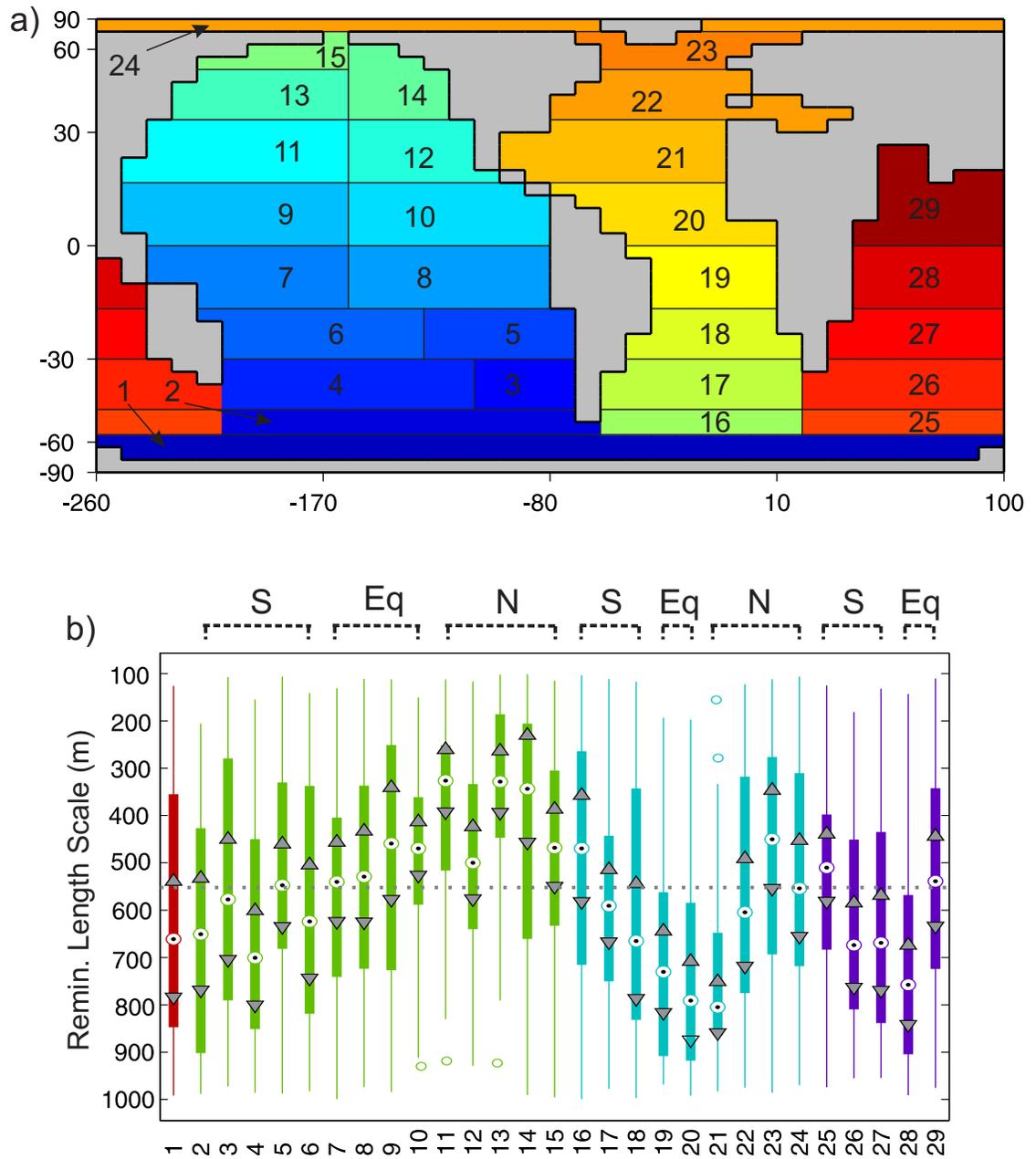
Figure C.8: Regions and optimal solution for n=29 regions. (a) Locations of regions used (b) boxplot of the 40 best solutions for each region comparable to Figure 5.10. 580 Latin hypercube samples were taken, of which 354 had a RMSD smaller than the CTRL run.