# CHRISTIAN-ALRECHTS-UNIVERSITÄT ZU KIEL

DISSERTATION

# Implications of Iron Model Complexity for the Projection of Global Biogeochemical Cycles

FROM THE PREINDUSTRIAL ERA AND INTO THE FUTURE

Dissertation zur Erlangung des Doktorgrades der Mathematisch-Naturwissenschaftlichen Fakultät der Christian-Albrechts-Universität zu Kiel vorgelegt von

> WANXUAN YAO Kiel 2021

Erster Gutachter und Betreuer: Prof. Dr. Andreas OSCHLIES Zweiter Gutachter und Betreuer: Prof. Dr. Eric ACHTERBERG Tag der Disputation: 23. Juni 2021

gez. Prof. Dr. Wolfgang Duschl, Dekan

# Contents

Lis	st of Tables	iii
Lis	st of Figures	v
Ac	ronyms	xi
0.	Summary	1
1.	Introduction1.1. Role of iron in global biogeochemistry1.2. The marine iron cycle1.3. Performance of iron models1.4. Parameter calibration difficulties and possible solutions1.5. University of Victoria Earth System Climate Model1.6. Chapter synopsis and author contribution	5 8 9 10 13 15
2.	Parameter sensitivity in the Kiel Marine Biogeochemical Model2.1. Introduction2.2. Methods2.3. Results and discussion2.4. Conclusions	<b>17</b> 17 18 23 35
3.	Hierarchy of calibrated global models reveals improved distributions and fluxes of biogeochemical tracers in models with explicit repre- sentation of iron3.1. Abstract	<b>37</b> 37 37 39 41 49
4.	Simulated future trends in marine nitrogen fixation are sensitive to model iron implementation4.1. abstract4.2. Introduction4.3. Method4.4. Results and Discussion4.5. Conclusions	<b>51</b> 51 53 55 62

5.	Conclusions   5.1. Outlook	<b>65</b> 67
bił	oliography	88
Α.	Supporting Information for 'Hierarchy of calibrated global models reveals improved distributions and fluxes of biogeochemical tracers in models with explicit representation of iron' A.1. Calibration framework	<b>89</b> 89 90
В.	Supporting Information for 'Simulated future trends in marine nitro- gen fixation are sensitive to model iron implementation' B.1. Nutrient and light limitation on phytoplankton growth	<b>97</b> 97
C.	Acknowledgements	101
D.	Declaration	103

# List of Tables

2.1.	The biogeochemical parameter values for Kiel Marine Biogeochemical Model Version 2 (KMBM2), summarized from Nickelsen et al. (2015)	19
2.2.	List of parameters in the sensitivity study and their low and high-cast values. The units of parameters are identical as in Tab. 2.1. For details about parameter values and references, please see the Results section below.	21
3.1.	The parameters chosen for each calibration, their observational value and calibration boundary settings. The parameters chosen for each calibration are marked with $\sqrt{2}$ .	41
3.2.	A comparison of the misfit from optimized model spin-ups. Values for uncalibrated FeDyn (FeDyn0) are calculated diagnostically (by us) and have not been used in the tuning process of Nickelsen et al. (2015). The	10
3.3.	minimum value of each column is marked in bold	42
A.1.	The parameters chosen for each calibration, their observational value and	чJ
ת 1	calibration boundary settings.	96
В.І.	Yao et al. (2019)	100

## List of Figures

1.1.	.1. Observed dissolved iron concentrations in the surface ocean (tiles; $\mu mol m^{-3}$ )				
	and locations that are primarily iron limited (red points within black cir-				
	cles). The iron concentrations are compiled from both the 2017 GEO-				
	TRACES intermediate data product (Schlitzer et al., 2018) and (Tagliabue				
	et al., 2012). The iron limitation locations are from Moore et al. (2013).	6			
1.0					

1.2. High nutrient low chloropyll regions: the North Pacific Ocean, the Equatorial Pacific Ocean, and the Southern Ocean. This is a screenshot of the video abstract by Yao et al. (2019) from link: https://iopscience.iop.org/article/10.1088/1748-9326/ab4c52.....

1.3.	The iron cycle in the ocean. This is a representation of the marine iron
	cycle adapted from the review on marine iron cycling by Tagliabue et al.
	(2017). The sources of iron in the ocean are aerosol deposition, sedimen-
	tary release, iceberg transport, and hydrothermal flux. Phytoplankton take
	up dissolved iron (dFe) and produce organic matter. Dissolved iron can be
	scavenged by particles, both organic and lithogenic, which form particu-
	late iron (pFe). Eventually particulate iron sinks out of the water column.
	However, not all dissolved iron will be scavenged. Ligands can be pro-
	duced as result of biological activity in the marine ecosystem and they
	can bind with iron and protect iron from scavenging. The ligand-bound
	iron can be then transported further by the ocean circulation (marked by
	dashed white arrow).
1.4.	Three different implementations of iron modules in a biogeochemistry
	model. The complexity increases from left to right (NoFe $<$ FeMask $<$
	FeDyn)
2.1.	Sensitivity of phytoplankton physiological parameters for different model tracers, represented by squares ( $NO_3$ ), stars ( $PO_4$ ), triangles ( $O_2$ ), and di-
	amonds $(dFe)$ Plot (a) is for the growth rate constant at $0^{\circ}C$ plot (b) is

24

7

- Sensitivity of phytoplankton light limitation parameters. The different 2.2. model tracers are represented by squares  $(NO_3)$ , stars  $(PO_4)$ , triangles  $(O_2)$ , and diamonds (dFe). Plot (a) is for the light attenuation by phytoplankton and plot (b) is for the maximum initial slope of the PI curve. Parameter values are represented by the colors purple (low-cast), green (ref), and yellow (high-cast). The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal RMSE and the interval between two neighboring line is 0.1. 26 2.3. Sensitivity of phytoplankton macro-nutrient limitation parameters. The different model tracers are represented by squares  $(NO_3)$ , stars  $(PO_4)$ , triangles  $(O_2)$ , and diamonds (dFe). Plot (a) is for the half-saturation constant of N and plot (b) is for the ratio of P and N in phytoplankton. Parameter values are represented by the colors purple (low-cast), green (ref), and yellow (high-cast). The observational data are located at point (1,0)and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal RMSE and the interval between 27 Sensitivity of phytoplankton iron limitation parameters. The shapes square, 2.4. star, triangular and diamond denote NO<sub>3</sub>, PO<sub>4</sub>, O<sub>2</sub>, and dFe respectively. The color purple, green, and yellow represents the low-cast, ref, and highcast respectively. Plot (a) is for the maximum iron half-saturation constant of Fe for the ordinary phytoplankton, plot (b) is for the threshold biomass concentration for the ordinary phytoplankton, plot (c) is for the iron halfsaturation constant of Fe for the diazotroph and plot (d) is the ratio of Fe and N in phytoplankton. The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal RMSE and the interval between two neighboring line is 0.1. The observations are located at the point [1,1] and closer a model to this point, the better its performance. . . . . . . . 28

31

33

2.6. Sensitivity of detritus parameters. The shapes square, star, triangular and diamond denote  $NO_3$ ,  $PO_4$ ,  $O_2$ , and dFe respectively. The color purple, green, and yellow represents the low-cast, ref, and high-cast respectively. Plot (a) is for the detritus remineralization rate at  $0^{\circ}C$ , plot (b) is for the sinking speed of detritus at surface, plot (c) is for the depth dependent sinking speed slope and plot (d) is the carbon-to-nitrogen ratio for organic matter. The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal *RMSE* and the interval between two neighboring line is 0.1.

2.7. Sensitivity of iron cycle parameters. The shapes square, star, triangular and diamond denote  $NO_3$ ,  $PO_4$ ,  $O_2$ , and dFe respectively. The color purple, green, and yellow represents the low-cast, ref, and high-cast respectively. Plot (a) is for the global ligand concentration, plot (b) is for the organic-matter dependent scavenging rate, plot (c) is for the organic-matter generation rate, plot (d) is the iron colloid production and precipitation rate, plot (e) is for the ratio between sedimentary iron release and organic phosphate, and plot (f) is for the minimum  $O_2$  concentration for aerobic oxidation. The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal *RMSE* and the interval between two neighboring line is 0.1.

ir is ), ie r- i) i- s, n of 3, x- as ie nt 42
d )- )) d es 11 46
1- 1, 1- ne 48
ı- s: 
1- 25 2) 1- 57
ts e: it ie te 58

4.4.	Change in annual export production at 130 <i>m</i> depth (years 2100 minus 1800; from left to right, NoFe, FeMask, FeDyn). Unit for the color scale is $gCm^{-2}yr^{-1}$ . The solid thick contour line denotes 0, and the solid (dashed) contour denotes positive (negative), the interval is $8 gCm^{-2}yr^{-1}$ . The dashed red rectangles represent the Eastern Tropical Pacific, Atlantic and Bay of Bengal regions. Note: the Caribbean region is excluded from the calculations for the box in the Eastern Tropical Pacific	60
4.5.	The change of annually averaged P* concentration $(mmol m^{-3})$ at 350 m depth between years 1800 and 2100. The solid thick contour line denotes 0, and the solid (dashed) contour denotes positive (negative) values, the interval of contour is 0.1 $mmol m^{-3}$ . While PO <sub>4</sub> and NO <sub>3</sub> were used as data constraints in the model calibration study (Yao et al., 2019), P* itself was not.	60
4.6.	A comparison of two different model responses to tropical warming and water column stratification, with (right side) and without (left side) explicit consideration of iron limitation. The upwelling ecosystem is controlled by a combination of top-down grazing and light limitation when iron does not limit primary production (left side). As the temperature increases, upwelling is reduced ("-" in the center of the figure). In the scenario not considering iron (left part of the figure) primary production increases faster than grazing pressure, which leads to increased particle export (+) and higher remineralization (+). Higher temperatures and increased remineralization intensify water column ODZs (+) and associated denitrification (+). Elevated denitrification leads to higher N <sub>2</sub> fixation, and an increase in diazotroph primary production downstream (+). However, including iron limitation (right side) mitigates the export production declines (-), and the ODZ volume also shrinks (-). More oxygen in the water column leads to less denitrification (-) and reduced P <sup>*</sup> , which suppresses N <sub>2</sub> fixation downstream despite temperature-driven increases in diazotroph primary production.	62
A.1.	Comparison between the annual basin-wide averaged nutrient and oxygen profiles in different basins from different calibrated models. Differences between the calibrated models are generally smaller than the differences between the models and observations. For example, in the Atlantic basin, the differences between calibrated models are much smaller (around 1 $mmol m^{-3}$ for $NO_3$ , 0.1 $mmol m^{-3}$ for $PO_4$ and 5 $mmol m^{-3}$ for $O_2$ ) than the differences between models and observations (around 6.5 $mmol m^{-3}$ for $NO_3$ , 0.6 $mmol m^{-3}$ for $PO_4$ and 20 $mmol m^{-3}$ for $O_2$ ). The similarities between the profiles of the three calibrated models suggest that they have approached some optimal limit for the given calibration setup. Addi-	

92

A.2.	Parameter trajectories in calibration of NoFe. The dashed straight lines denote the parameter calibration boundaries and the solid straight line de-					
	notes the uncalibrated parameter value. The bold black line is the average					
	parameter value in each generation and the thin red lines shows the max-					
	imum and minimum value parameter value in each generation	93				
A.3.	Parameter trajectories in calibration of FeMask. The notation follows fig-	04				
		94				
A.4.	Parameter trajectories in calibration of FeDyn. The notation follows fig-					
	ure A.3	95				
R 1	The most limiting growth factor for ordinary phytoplankton in all models					
	The most minume growth factor for ordinary phytophankton in an models					
2.11.	in the years 1800 (unner) and 2100 (lower)	90				
D.1.	in the years 1800 (upper) and 2100 (lower).	98				
B.2.	in the years 1800 (upper) and 2100 (lower)	98				
B.2.	in the years 1800 (upper) and 2100 (lower)	98				
B.2.	in the years 1800 (upper) and 2100 (lower)	98				
B.2.	in the years 1800 (upper) and 2100 (lower)	98				
B.2.	in the years 1800 (upper) and 2100 (lower)	98				
B.2.	in the years 1800 (upper) and 2100 (lower)	98				
B.2.	in the years 1800 (upper) and 2100 (lower)	98 99				

#### Acronyms

- **CMA-ES** Covariance Matrix Adaption Evolution Strategy
- **EDA** Estimation of Distribution Algorithm
- EMIC Earth system Model of Intermediate Complexity
- **ETP** Eastern Tropical Pacific
- FCT second-order Flux-Corrected Transport scheme
- GCM ocean General Circulation Models
- HNLC High Nutrient Low Chlorophyll
- KMBM Kiel Marine Biogeochemical Model
- KMBM1 Kiel Marine Biogeochemical Model Version 1
- KMBM2 Kiel Marine Biogeochemical Model Version 2
- MOM2 Modular Ocean Model 2
- **NPP** Net Primary Production
- NPZD Nutrient-Phytoplankton-Zooplankton-Detritus
- **ODZ** Oxygen Deficient Zone
- OMZ Oxygen Minimum Zone, traditional name for ODZ
- **RCP** Representative Concentration Pathway
- **RMSE** Root Square Mean Error
- **TM** Transport Matrix
- **TMM** Transport Matrix Method
- UVic 2.9 University of Victoria Earth System Climate Model version 2.9
- UVic ESCM University of Victoria Earth System Climate Model
- UW3 Upwind-biased scheme

## 0. Summary

The trace metal iron is considered to be the nutrient that limits marine primary production in one third of the global surface ocean (Martin, 1990; Boyd et al., 2007; Moore et al., 2013). It is also the nutrient that maintains future ocean fertility due to its irreplaceable role in the process of nitrogen fixation, which adds "new" nitrogen (another nutrient for phytoplankton) to the surface ocean (Raven, 1988; Kustka et al., 2003b; Zehr and Capone, 2020).

Due to iron's importance, it is not surprising that the demand for incorporating iron into global biogeochemical models is high. However, including iron in an earth system model has been shown to have no clear benefits with respect to model misfit against observational data (Nickelsen et al., 2015). How smart is it then to introduce iron models into global biogeochemical models, when the benefits are not clearly identifiable? Especially, when the iron models perform poorly at reproducing observed iron patterns in the ocean (Tagliabue et al., 2016).

The poor performance of iron models, coupled with their failure to improve biogeochemical tracer representation of the ocean, inspired this additional effort to identify the advantages of including iron in a global biogeochemical model, both for the preindustrial state and under conditions of a changing climate. The working hypothesis was that the relatively poor performance of iron models might come from inadequate model calibration.

A first sensitivity study on biogeochemical model parameter values was conducted in order to identify key parameters for model calibration. It was found that while some of the parameters influence simulated nitrogen, phosphorus, and oxygen concentrations, few parameters influence simulated iron concentrations. This suggests that our modelling skill of the iron cycle is still limited and/or that the observational data base is insufficient for comprehensive model calibration so far. Thus it was decided not to include iron data in further model calibration.

A model calibration framework (Kriest et al., 2017) was next applied to a hierarchy of global models with different implementations of iron; one without iron, one with prescribed iron concentrations, and another one with a dynamic iron cycle. Using calibration against global data sets of nitrogen, phosphorus, and oxygen, the misfit of each model was pushed to its minimum. It was found that under an assumed preindustrial steady state, the calibrated model with a full dynamic iron cycle has the lowest model misfit against observations (thus confirming the working hypothesis). It was also found that the calibrated model with a fully dynamic iron cycle has 50% less net primary production (which is closer to empirical estimations) compared to the calibrated model without iron.

Finally, transient simulations for all calibrated models were integrated from their preindustrial state until the end of the 21st century using an atmospheric  $CO_2$  concentration pathway consistent with a 'business-as-usual'  $CO_2$  emission scenario. It was found that nitrogen fixation trends diverge among models. This divergence is caused by whether iron limits the productivity of the upwelling regions, e.g. in the eastern tropical Pacific. The export production in the eastern tropical Pacific (and other tropical upwelling regions) reacts differently to warming, depending on whether iron is a limiting nutrient. These different responses trigger a divergent chain of downstream responses that affect nitrogen fixation across the tropical oligotrophic regions in the model.

Through the comparison between calibrated models, this thesis quantifies the advantages of including iron in a global biogeochemistry model and reveals how important iron is for future nitrogen fixation trends. It furthermore illustrates the interconnection between tropical upwelling and oligotrophic regions.

## Zusamenfassung

Das Spurenmetall Eisen gilt als der Nährstoff, der die Primärproduktion in etwa einem Drittel der globalen Ozeanoberfläche begrenzt (Martin, 1990; Boyd et al., 2007; Moore et al., 2013). Da Eisen bei der Stickstofffixierung nicht ersetzt werden kann, durch die dem Oberflächenozean neuerSStickstoff (ein weiterer Nährstoff für Phytoplankton) zugeführt wird, hält Eisen die Ozean-Fertilität aufrecht (Raven, 1988; Kustka et al., 2003b; Zehr and Capone, 2020).

Angesichts der Bedeutung von Eisen ist es nicht überraschend, dass der Bedarf einer Implementierung des ozeanischen Eisenkreislaufs in globale biogeochemische Modelle hoch ist. Die Einbeziehung von Eisen in Erdsystemmodellen hat jedoch bislang keine eindeutigen Vorteile hinsichtlich des Modell-Misfits gegenüber Beobachtungsdaten gezeitigt (Nickelsen et al., 2015). Wie sinnvoll ist es dann, Eisenmodelle in globale biogeochemische Modelle zu implementieren, wenn der Nutzen nicht klar erkennbar ist? Vor allem, wenn die Eisenmodelle die beobachteten Muster im Ozean nur unzureichend reproduzieren (Tagliabue et al., 2016).

Die schwache Leistung bisheriger Eisenmodelle, einhergehend mit deren Unfähigkeit, die simulierte Verteilung biogeochemischer Tracer im Ozean zu verbessern, inspirierte diese zusätzliche Forschung, um zu eruieren, welche Vorteile die Einbeziehung von Eisen in ein globales biogeochemisches Modell haben könnte, sowohl für den vorindustriellen Zustand als auch unter den Bedingungen eines sich ändernden Klimas. Die Arbeitshypothese war, dass die relativ schwache Leistung der Eisenmodelle auf eine unzureichende Modellkalibrierung zurückzuführen sein könnte.

Um die Schlüsselparameter für die Kalibrierung zu identifizieren, wurde eine erste Sensitivitätsanalyse der biogeochemischen Parameter durchgeführt. Die Ergebnisse der Studie zeigen, dass einige Parameter zwar die simulierten Stickstoff-, Phosphor-, und Sauerstoffkonzentrationen beeinflussen, aber nur wenige Parameter die Eisenkonzent-rationen. Dies weist darauf hin, dass unsere Fähigkeit, den Eisenkreislaufs zu simulieren, noch begrenzt ist, bzw. dass die Datenbasis für eine umfassende Modellkalibrierung bisher unzureichend ist. Daher wurden die Eisendaten nicht in die weitere Modellkalibrierung einbezogen.

Im nächsten Schritt wurde eine Hierarchie globaler Modelle mit unterschiedlichen Implementierungen von Eisen in einem Modell-Kalibrierungs-Framework (Kriest et al., 2017) analysiert. Eines der Modelle ist ohne Eisen, eines hat vorgeschriebene Eisenkonzentrationen und ein weiteres einen dynamischen Eisenkreislauf. Die Kalibrierung hat den Misfit der Modelle gegen globale Datensätze von Stickstoff, Phosphor und Sauerstoff minimiert. Unter der Annahme eines vorindustriellen Steady States hatte das kalibrierte Modell mit einem dynamischen Eisenzyklus den geringsten Modell-Missfit gegenüber den Beobachtungen, wodurch die Arbeitshypothese bestätigt wurde. Außerdem weist das kalibrierte Modell mit dynamischem Eisenzyklus 50 % weniger Netto-Primärproduktion auf (was näher an empirischen Schätzungen liegt) als das ohne Eisen.

Abschließend wurden transiente Simulationen für alle kalibrierten Modelle vom vorindustriellen Zustand bis zum Ende des 21. Jahrhunderts integriert, mittels eines atmosphärischen CO<sub>2</sub>-Konzentrationspfades, der mit einem "Business-as-usual"-CO<sub>2</sub>-Emissionsszenario konsistent ist. Die Trends der Stickstofffixierung divergieren zwischen den Modellen. Die Divergenz wird dadurch verursacht, ob Eisen die Produktivität der Auftriebsgebiete, z.B. im östlichen tropischen Pazifik, begrenzt. Die Exportproduktion im östlichen tropischen Pazifik (und anderen tropischen Auftriebsgebieten) reagiert unterschiedlich auf die Erwärmung, je nachdem, ob Eisen ein limitierender Nährstoff ist. Diese unterschiedlichen Reaktionen lösen eine divergente Kette von nachgelagerten Reaktionen aus, die die Stickstofffixierung in den tropischen oligotrophen Regionen der Modelle beeinflussen.

Durch den Vergleich zwischen kalibrierten Modellen quantifiziert diese Dissertation die Vorteile der Einbeziehung von Eisen in ein globales biogeochemisches Modell und zeigt, wie wichtig Eisen für zukünftige Trends der Stickstofffixierung ist. Zudem veranschaulicht diese Dissertation die Verbindung zwischen tropischen Auftriebsgebieten und oligotrophen Regionen.

## 1. Introduction

#### 1.1. Role of iron in global biogeochemistry

Tiny phytoplankton form the base of the marine ecosystem and help to shape the chemical and elemental distribution of the Earth's vast oceans (Redfield, 1958). The marine ecosystem provides essential services, such as fisheries and recreation, for human society. In a photosynthetic process called "primary production", phytoplankton take up dissolved carbon dioxide and form organic material, some of which passes through the surface ecosystem and eventually sinks into the deep ocean where it decomposes back to CO<sub>2</sub>. This sequence of processes, often referred to as the biological carbon pump, is an important component of the marine carbon cycle. Its carbon uptake potential is known to play a considerable role in shaping Earth's climate on long time scales (Sarmiento and Gruber, 2006) and might also play a role in the oceanic sequestration of carbon as we face the challenge of climate change in the anthropocene (Riebesell et al., 2009; Shaffer, 2010; Passow and Carlson, 2012; Taucher et al., 2014).

However, the nutrients required for phytoplankton growth are not always available in the surface ocean and therefore primary production can be nutrient-limited (Moore et al., 2013). Over the years, there has been debate over whether nitrogen or phosphorus is the ultimate limiting nutrient in the ocean (Tyrrell, 1999). A "geologist's view" is that phosphorus is the ultimate nutrient limiting global primary production. This is based on the existence of a process called 'nitrogen fixation'. Nitrogen fixation occurs in nitrogendepleted waters, where a special type of phytoplankton, diazotrophs, fix dinitrogen and make it bio-available. Phosphorus does not have an analogous route of production within the ocean; hence, phosphorus can eventually become more scarce compared to nitrogen, e.g. if loss by burial of organic matter in sediments exceeds external terrestrial sources of phosphate globally. Furthermore, a "biologist's view" suggests that nitrogen is the ultimate limiting nutrient. This is indicated by the ratio of nitrogen to phosphorus in the surface water showing a relative shortage of nitrogen in observational data, e.g., World Ocean Atlas (Tyrrell and Law, 1997). Additionally, it has been shown that adding nitrogen to oligotrophic, or nutrient-poor, waters can induce higher growth while adding phosphorus does not (Ryther and Dunstan, 1971; Moore et al., 2013). Tyrrell (1999) used an ocean box model to explain that in anoxic regions, denitrification acts as a nitrogen sink, contributing to the shortage of nitrogen in the surface. They found that denitrification is globally and on long timescales balanced by nitrogen fixation and both are controlled by the addition of phosphorus. Hence, they concluded that phosphorus is the ultimate limiting nutrient regulating long-term ocean productivity. However, Tyrrell (1999) did not consider the constraints on nitrogen fixation in their model, e.g. by iron availability. The availability of iron defines the biogeography of diazotrophs in the low latitudes and is thus important for the long-term stability of the nitrogen inventory, and the maintenance of ocean fertility (Ward et al., 2013; Snow et al., 2015).



**Figure 1.1.:** Observed dissolved iron concentrations in the surface ocean (tiles;  $\mu mol m^{-3}$ ) and locations that are primarily iron limited (red points within black circles). The iron concentrations are compiled from both the 2017 GEOTRACES intermediate data product (Schlitzer et al., 2018) and (Tagliabue et al., 2012). The iron limitation locations are from Moore et al. (2013).

Iron as a micro-nutrient is irreplaceable in certain cellular enzymes associated with photosynthesis, respiration, and nitrogen fixation (Raven, 1988; Kustka et al., 2003b; Tagliabue et al., 2017; Zehr and Capone, 2020), owing to the origin of marine life in iron rich anoxic conditions (Canfield et al., 2006). In contrast to its abundance in the Earth's crust, iron is not among the most abundant chemical compounds (e.g. chloride, sodium, sulfate, magnesium, calcium, and potassium) present in modern seawater. This is due to the low solubility of iron in oxygenated water and a process called "scavenging", where dissolved iron attaches itself to particles and sinks through the water column. The concentration of iron in seawater is on the scale of  $\mu mol m^{-3}$  or lower (Fig. 1.1), which is about a thousand times lower than the molar concentration of phosphate. This concentration is so low that it was not until the 1970s, with the achievement of contamination-free sampling and trace-metal measurements (Bruland et al., 1979; Settle and Patterson, 1980), that it was possible to accurately measure iron concentrations in seawater. Considering that the stoichiometric ratio between iron and phosphate in phytoplankton varies from 4.6 to 31 mmol mol<sup>-1</sup> (Twining and Baines, 2013) and the ratio in seawater is about 0.27  $mmol \ mol^{-1}$  (Moore et al., 2013), it is not surprising that iron in seawater is a rare com-



**Figure 1.2.:** High nutrient low chloropyll regions: the North Pacific Ocean, the Equatorial Pacific Ocean, and the Southern Ocean. This is a screenshot of the video abstract by Yao et al. (2019) from link: https://iopscience.iop.org/article/10.1088/1748-9326/ab4c52.

modity for phytoplankton.

Primary production in one-third of the surface ocean is limited by iron and the shortage of iron explains the existence of "High-Nutrient-Low-Chlorophyll" (HNLC) regions, such as the Southern Ocean, where relatively high concentrations of nitrogen and phosphorus in the surface water are not utilized by biological production (Fig. 1.2; Boyd et al., 2007; Moore et al., 2013). Based on these findings, could iron be the actual ultimate limiting nutrient of primary production? The answer is complicated. The in-situ iron fertilization experiments show an increase of the primary production during the iron addition period (de Baar et al., 2005). The global iron fertilization based on model studies show that the primary production and carbon uptake of the ocean may increase (Aumont and Bopp, 2006; Zahariev et al., 2008). However, in those models experiments, simulating replete iron for a short period, the phytoplankton community in the HNLC regions can utilize more micro-nutrients causing in turn a reduction of macro-nutrient lateral transport to oligotrophic regions and a decrease of productivity downstream.

Martin (1990) hypothesized that the lower atmospheric  $CO_2$  concentrations of the Last Glacial Maximum (LGM) may have been due to a much higher ocean carbon export from enhanced surface biological activity. They hypothesized that this enhanced carbon export was caused by an increase in iron supply due to higher levels of atmospheric dust deposition. This hypothesis, famously described by Martin with his controversial statement,

"Give me half a tanker of iron, and I'll give you the next ice age", has inspired multiple mesoscale iron fertilization experiments (de Baar et al., 2005) and ideas for climate engineering (Aumont and Bopp, 2006). Model studies show that iron fertilization might not be the sole factor for the carbon dioxide drop during the LGM, but enhanced productivity in high latitudes might have contributed from less than a quarter (Lambert et al., 2015) to one-half (Watson et al., 2000; Hain et al., 2010) of the decrease in atmospheric carbon dioxide. This uncertainty, in addition to the circulation and forcing differences, could be partly caused by differences between model implementations (including model complexity) of the iron cycle (Tagliabue et al., 2017).

### 1.2. The marine iron cycle



**Figure 1.3.:** The iron cycle in the ocean. This is a representation of the marine iron cycle adapted from the review on marine iron cycling by Tagliabue et al. (2017). The sources of iron in the ocean are aerosol deposition, sedimentary release, iceberg transport, and hydrothermal flux. Phytoplankton take up dissolved iron (dFe) and produce organic matter. Dissolved iron can be scavenged by particles, both organic and lithogenic, which form particulate iron (pFe). Eventually particulate iron sinks out of the water column. However, not all dissolved iron will be scavenged. Ligands can be produced as result of biological activity in the marine ecosystem and they can bind with iron and protect iron from scavenging. The ligand-bound iron can be then transported further by the ocean circulation (marked by dashed white arrow).

Our understanding of the iron cycle has evolved in the last few decades. Before the GEO-TRACES project (Schlitzer et al., 2018) discovered variable iron concentrations in the deep ocean linked to hydrothermal vents (e.g., Fitzsimmons et al., 2015), the deep ocean iron concentrations were assumed to be quasi-constant and reflecting the ambient organic iron-complexing ligand concentrations (Gledhill and van den Berg, 1994; Rue and Bruland, 1995). Aerosols were assumed to be the sole source of iron in the ocean in early model studies (Lefèvre and Watson, 1999; Archer and Johnson, 2000). This was partly motivated by earlier studies that suggested a limited reach of riverine iron; e.g., rapid removal of iron in river water by salt-induced flocculation in estuary regions (Boyle et al., 1977). More recently, Elrod et al. (2004) found out that the shelf sediments can be a significant iron source and can sustain high phytoplankton productivity downstream hundreds of kilometers offshore. Robinson et al. (2016) also pointed out that there is natural iron fertilization downstream from islands, primarily due to sediments and runoff. In the polar oceans, icebergs can be a major iron source, since the iron particles can hitchhike on the iceberg during its voyage (Raiswell et al., 2008; Hopwood et al., 2019). Although hydrothermal sources are important for the deep ocean iron inventory, they are a minor contributor to ocean primary production, fertilizing only 2 to 3 percent of the global ocean surface production of particle organic carbon and is less pronounced than aerosol (around 12 to 15%) and sediment sources (79 to 81%; Tagliabue et al., 2014).

In the vast ocean, the rapid removal of iron by scavenging process seems to set a general trend of iron concentration: low abundance when local iron source is absent. However, organic ligands that bind with iron can keep iron in solution (Fig. 1.3). The production of ligands is associated with iron-limited bacterial and phytoplankton community growth, zooplankton grazing and organic particle remineralization (Gledhill, 2012). Völker and Tagliabue (2015) introduced a dynamic ligand cycle in a global biogeochemical model and found improvement for model misfit against observational iron patterns.

#### 1.3. Performance of iron models

Marine biogeochemical models are a means to test hypotheses induced by observations or experimental work, to illustrate interactions between marine ecosystems and physical environments, and to deduce the implications of such interactions. As more processes are discovered and hypotheses are generated, for example by novel observational data (e.g., the finding of hydrothermalism being an iron source in the interior ocean; Tagliabue et al., 2010), there is increasing demand to include these proposed processes into biogeochemical models. This demand is based on the assumptions that (a) a more complex model better reflects reality and (b) that a model that better reflects current observations provides for more realistic future projections. Neither of which are necessarily the case. Modelers can be hesitant to implement more processes into models, or to make models more complex. This "hesitation", however, may not be about the implementation process itself but rather about the challenges that come with the calibration of more complex models.

Increasing in complexity without sufficient parameter calibration does not guarantee better model projections (Anderson, 2005). In an iron model inter-comparison project against observed data of dissolved iron by Tagliabue et al. (2016), the authors compared 13 global biogeochemical models and found large disagreements in iron residence time in the ocean (from the scale of years to centuries). This variability is due to differences in the applied iron scavenging parameterizations. The authors also pointed out that even the most complex iron model has a large model-data misfit (and only explained 36 % of the observed iron concentration variance, despite representing a wide variety of processes thought to be relevant). Nickelsen et al. (2015) hand-tuned parameter values against surface nutrient fields as they introduced a full dynamic cycle of iron into an earth system model with intermediate complexity (Eby et al., 2013; Keller et al., 2012). This addition of iron did not lead to improvement of model performance with respect to reproducing observed surface nutrient distributions, compared to an earlier version of the biogeochemical model by Keller et al. (2012). This earlier version uses a simple prescribed iron concentration mask in the surface ocean to limit primary production. Poor parameter value choices in the new, more complex model might have contributed to the stagnating performance, but also this might not be the sole issue. In addition, the uncertainty in iron models needs to be addressed before we can trust model results that examine iron fertilization as a potential climate change mitigation measure (Tagliabue et al., 2016; Buesseler et al., 2008). More effort should be put into improving parameter calibration techniques, and into understanding the impact of model complexity before increasing model complexity further.

Parameter uncertainty may occur not in iron modules alone, but is widespread in the global biogeochemical models. Model architecture and its parameters describe the pathways for elements (e.g., nutrients and carbon) passing through marine ecosystems. Even when models show similar nutrient distributions (e.g., nitrate, phosphate), the fluxes in the ecosystem can vary (Löptien and Dietze, 2017). Laufkötter et al. (2016) show that the projections of carbon export over the 21st century vary across different models due to pathway uncertainties and call for better-constrained ecosystem parameterizations.

#### 1.4. Parameter calibration difficulties and possible solutions

In this section I will address the question, what is hindering calibration of model parameters in an earth system model.

#### 1.4.1. Computing expense of a single simulation

In coupled physical-biogeochemical models of the ocean, the global chemical and biological tracers (e.g., the concentration of dissolved inorganic nitrogen and the abundance of phytoplankton) are generally modeled within the framework of local "grid-boxes", which are connected via the ocean General Circulation Models (GCM). This means that the chemical and biological processes take place in the local grid-box and the physical transport (advection and mixing) of tracers between the grid-boxes is carried out by the GCMs. Since many tracers in the ocean model require thousands of model years to reach their equilibrium state (i.e., the annual mean concentration of a tracer does not change over time), the effect of changes in parameter values on steady state solutions of ocean models are computationally expensive to obtain. A fully coupled earth system model used for assessing the impact of climate change in the Climate Model Inter-comparison Projects (CMIP; e.g., Bopp et al., 2013; Eyring et al., 2016) might take up to a month to finish a 300-year simulation. Due to this long integration time, even a single spin-up (the simulation where a model reaches its equilibrium state) may take up to a year.

Khatiwala (2007) reduced the model spin-up time problem with the Transport Matrix Method (TMM). The underlying idea of the TMM is that a GCM, if the advectivediffusive transport operator is linear, it can be written as a sparse matrix, which may be efficiently constructed by "probing" the GCM with a passive tracer (Khatiwala, 2007). This sparse matrix is called a Transport Matrix (TM). The TM is used to move tracers around without activating its corresponding GCM: hence, it is an "offline model". Furthermore, models using a seasonally cycling set of constant TM by Khatiwala (2007) are adapted to parallel computing frameworks. Consequently, the model spin-up can be run at up to two orders of magnitude higher efficiency. However, this offline model has some bias to biogeochemical quantities (e.g., differences in tracer concentrations at high latitudes) compared to its online counterpart. This is due to modifications (e.g., omission of the polar filtering applied in the online version, and slight differences in the application of time-dependent forcing fields) that are necessary in the TMM (Kvale et al., 2017).

#### 1.4.2. Dealing with large numbers of parameters

Model complexity increases and the number of tunable parameters rise. Given the available observations, in particular the number of independent datasets, not all parameters can be independently tuned in biogeochemical models (Kriest, 2017). Searching a combination of parameter values that minimize misfit between model results and observational data in a model with *n* parameters is a search problem in a solution space with *n* dimensions (e.g., if there are two value choices for each parameter, a model with 46 parameters has  $2^{46}$  possible parameter value combinations). This means that when *n* increases by 1, the number of combinations that need to be tested increases at least by a twofold. Hence, hand-tuning or making educated guesses, can be ever more challenging when dealing with exponentially expanding numbers of possible parameter value combinations.

One of the possible solutions for parameter optimization is the Covariance Matrix Adaption Evolution Strategy (CMA-ES) (Kriest et al., 2017; Hansen, 2006). It is one of the meta-heuristic methods, which is made to solve the minimization problem in multiple dimensions by mimicking the evolution process in nature to search the parameter space efficiently. For the classical evolutionary algorithms, parameter value is seen as a trait and a combination of parameter values is seen as an individual. The performance of an individual is measured by a cost function (e.g., the misfit against observations of the model which applied this very combination of parameter values) and a "better" individual is an individual with a lower cost (i.e., a smaller misfit). In the beginning, a certain amount of individuals are "born" by drawing randomly from the solution space (i.e. a plausible range of values for a given parameter as indicated by e.g. the respective physiological literature). In each generation or iteration, the performance (misfit) of individuals will be evaluated and the offspring will be produced by random selection of traits from the "better" individuals, while "worse" individuals will not be considered further (they go 'extinct'). The algorithm will also allow the re-insertion of "good" individuals from former generations in order to give the good traits (parameter values) a better chance to pass on to the next generation of individuals. This evolutionary process encourages the emergence of individuals with preferable traits. After a certain number of generations, the "best" individual with the smallest misfit arises.

Different from classical evolution algorithms, CMA-ES is an Estimation of Distribution Algorithm (EDA) (Hansen and Ostermeier, 2001). CMA-ES assesses the distribution of each parameter not only from the misfit of the current generation but also a few generations backwards, and looks for where is more likely to have better fitness. In this way, the knowledge of the distribution gathered by older generations can persist and fade out slowly. CMA-ES then draws the offspring following the knowledge of the distribution. The advantage of CMA-ES is that it needs smaller amounts of individuals per generation, and hence lower computational demand compared to classical evolutionary algorithms. However, considering the total numbers of parameters in a biogeochemical model and the limited number of available independent data sets (see paragraph below), the values of some parameters may not converge after calibration (Löptien and Dietze, 2015, 2017). In a study by Kriest et al. (2017), a selective set of 6 parameters from a biogeochemical model (which contains over 20 parameters) is calibrated against the global data sets of nitrate, phosphate, and oxygen concentrations, and the calibrated model shown a better skill to reproduce observed global biogeochemical fluxes.

#### 1.4.3. Data set limitations

As the models grow in complexity, we need more independent data sets to constrain the parameters. The World Ocean Atlas is an precious asset for modellers. It contains a climatological compilation of a few biogeochemical tracers, such as oxygen, nitrate, phosphate, and silicate in the world ocean (Garcia et al., 2013b,a). However, the observations for iron in the ocean is relatively scarce, since, as motioned above, it was not possible to measure the low concentration of iron in sea water until the late 70s (Bruland et al., 1979; Settle and Patterson, 1980), and still requires dedicated equipment and experts. Thanks to in-

ternational collaborations, e.g., the GEOTRACES project (Schlitzer et al., 2018), limited amount of cruise sections of iron concentration data in all major basins are collected or planned to be collected. These can provide insight into the variability of iron concentrations in the ocean.

Observation-based biogeochemical flux/rate measurements or estimates (e.g. net primary production and export production) may provide extra constraints for biological parameters, in addition to traditional concentration-based data sets. However, fluxes have their own uncertainty. Carr et al. (2006) compared 24 models that estimate global ocean net primary production from satellite measurements of ocean color and found a variation by a factor of two among them, contributed possibly by different model estimates in high-nutrient low-chlorophyll conditions, and extreme temperatures or chlorophyll concentrations. Sediment traps are criticized for systematic bias (e.g., caused by tipping over, advection, sample preservation, and estimation models) (Honjo et al., 2008; Bloesch and Burns, 1980). They are still an important data set for assessing the global oceanic particle sinking rates, because of their wide deployments in various regions of the ocean ever since 1983 (Honjo et al., 2008).

#### 1.5. University of Victoria Earth System Climate Model

Earth system models of intermediate complexity (EMICs) include most of the processes described in comprehensive models and have a coarse grid (compare to higher resolution in comprehensive models) and often apply substantial simplifications to represent the dynamics of some or all components, which make the models simple enough that long-term climate simulations over thousands of years are sensible (weeks of computing) on the current generation of computers (Claussen et al., 2002). University of Victoria Earth System Climate Model version 2.9 (UVic 2.9) is an EMIC with four coupled components (Weaver et al., 2001; Eby et al., 2013): a single layer atmospheric model, a sea ice model (Bitz and Lipscomb, 1999), a land model (Meissner et al., 2003), and a three dimensional ocean model. In the horizontal direction, all components have a grid resolution of 1.8 ° latitude  $\times$  3.6 ° longitude. The atmospheric model calculates heat and water fluxes between the atmosphere and the ocean, land and sea ice dynamically (Fanning and Weaver, 1996). The vertical grid resolution of the ocean model varies from 50 *m* near the surface to 500 *m* in the deep ocean.

In order to isolate the impact on biogeochemistry from the complexity of the iron models, we apply the same GCM (Modular Ocean Model 2, Pacanowski, 1996), coupled with different versions of the Kiel Marine Biogeochemical Model (KMBM) as the oceanic biogeochemical component (Fig. 1.4). The first model ignores iron as a nutrient for phytoplankton (NoFe) and the phytoplankton potential growth rate depends on nitrate, phosphate, light intensity, and temperature. This version of the model is created by dis-



**Figure 1.4.:** Three different implementations of iron modules in a biogeochemistry model. The complexity increases from left to right (NoFe < FeMask < FeDyn).

abling the iron limitation factor in (KMBM1, Keller et al., 2012). The second model has a prescribed seasonal cycling of iron concentrations (FeMask, Keller et al., 2012), which is estimated by the Biology Light Iron Nutrient and Gas model (BLING, Galbraith et al., 2010). In FeMask, the iron limitation factor (a Monod function of iron concentrations) acts as an additional constraint for the phytoplankton potential growth rate compared to NoFe. The third model builds on FeMask and instead of using prescribed iron concentrations, it calculate iron concentrations dynamically through an explicit representation of the oceanic iron cycle (FeDyn, Nickelsen et al., 2015) similar to what was used originally in the BLING model simulations. In FeDyn, the accounts for different iron sources, such as aerosol deposition (Luo et al., 2008), sediment release (Elrod et al., 2004), and hydrothermal iron flux (Tagliabue et al., 2010). Dissolved iron in FeDyn is also subjected to scavenging and colloid formation. Ligands can bind with dissolved iron and prevent iron scavenging, which is prescribed as a global constant concentration in FeDyn. More details of each model variation can be found in Keller et al. (2012, NoFe and FeMask) and Nickelsen et al. (2015, FeDyn).

#### 1.6. Chapter synopsis and author contribution

This thesis aims to understand the impact of different iron model implementations on biogeochemical processes under pre-industrial steady state and climate change. The goals and achievements of each chapter are summarized here.

**Chapter 2** aims to identify suitable parameters (i.e. those to which simulated tracer distribution is sensitive) in UVic 2.9 for model calibration in Chapter 3. This chapter focuses on identifying parameter uncertainties from the literature and testing parameter sensitivity for the pre-industrial climate state in the offline UVic 2.9. It demonstrates possible deficiencies arising from hand-tuning and the major difficulties of the iron model at reproducing observational data. It paints a general picture for parameter sensitivities in UVic 2.9 with KMBM 2(Nickelsen et al., 2015). It also lays down some of the ground work for the model calibrations in the next chapter.

A. Oschlies, W. Koeve and W. Yao conceived and designed the experiments. W. Yao implemented and performed the experiments and analysed the data. W. Yao wrote the chapter with contributions from K. F. Kvale, W. Koeve and A. Oschlies.

**Chapter 3** aims to explore the impact of different iron implementations on global biogeochemical indicators in a pre-industrial climate after calibration, which optimizes every model against observational data. In this chapter a model calibration framework by Kriest et al. (2017) is applied that utilized the TMM and the CMA-ES. Three different variations of UVic 2.9 (one without iron implementation, one with prescribed iron concentrations, and another with a full iron cycle) are calibrated against multiple observational data fields (nitrogen, phosphate, and oxygen). By comparing the calibrated models, the impact of iron model implementations on global biogeochemical indicators (e.g., net primary production and oxygen deficit zone volume) in the pre-industrial steady state ocean are quantified. It is demonstrated that the model with a full iron cycle implementation has the best performance. Different nutrient pathways resulting from different iron model implementations are described, which emphasize the importance of model parameter calibration for future studies.

This chapter is based on the published paper: Yao W, Kvale K F, Achterberg E, Koeve W, Oschlies A. (2019) Hierarchy of calibrated global models reveals improved distributions and fluxes of biogeochemical tracers in models with explicit representation of iron[J]. Environ. Res. Lett. 14 114009, DOI:https://doi.org/10.1088/1748-9326/ab4c52. W. Yao and A. Oschlies conceived and designed the experiments. W. Yao implemented and performed the model calibrations with support from K. F. Kvale and W. Koeve. W. Yao analysed the data. W. Yao wrote the paper with contributions from K. F. Kvale, W. Koeve, E. Achterberg and A. Oschlies.

**Chapter 4** aims to examine the impact of applying different iron implementations under climate change. The three calibrated variations of UVic 2.9 (Chapter 3) are integrated online (in a fully-coupled earth system mode) under the atmospheric  $CO_2$  concentra-

tion scenario Representative Concentration Pathways (RCP) 8.5 of the Intergovernmental Panel on Climate Change (IPCC). While all models agree on a trend of decreasing NPP with global warming, models disagree on nitrogen fixation trends. Comparing all models, we find that the nutrient pathways in the upwelling regions, e.g., in the tropical eastern Pacific, are determined by model differences in the iron cycle parameterization. These pathways behave differently in response to warming, which creates different nutrient sto-ichiometric conditions (nitrogen to phosphorus ratio) for diazotrophs in the oligotrophic regions, and hence triggers different nitrogen fixation responses. This chapter illustrates the importance of a realistic iron parameterization for future climate studies and argues for more effort in model calibration accompanied with future model development. This chapter is based on a manuscript submitted to the scientific journal Global Biogeochemical Cycles: W. Yao, K. F. Kvale, W. Koeve, A. Landolfi, E. Achterberg, E. M. Bertrand and A. Oschlies (2020) Simulated future trends in marine nitrogen fixation

are sensitive to model iron implementation. W. Yao conceived and designed the experiments. W. Yao performed the model simulation and analysed the data. W. Yao wrote the manuscript with contributions from K. F. Kvale, W. Koeve, A. Landolfi, E. Achterberg, E. M. Bertrand and A. Oschlies.

# 2. Parameter sensitivity in the Kiel Marine Biogeochemical Model

#### 2.1. Introduction

Biogeochemical models are an important component of earth system models, and describe marine ecological interactions with chemical elements and the physical environment. Over the years, biogeochemical models have become more complex. For example, the KMBM originated as a simple nitrogen-based NPZD pelagic model (Oschlies and Garçon, 1999), which contained only 4 biological tracers; nitrogen (N), phytoplankton (P), zooplankton (Z) and detritus (Det) and only 16 biogeochemical parameters. Diazotrophs were then added by Schmittner et al. (2008); these fix dinitrogen when fixed (plant usable) nitrogen in the water is scarce. In the Kiel Marine Biogeochemical Model Version 1 (KMBM1), Keller et al. (2012) implemented a global mask of iron-related growth rate limitation for phytoplankton, based on the mean monthly dissolved iron concentration outputs from the BLING model (Galbraith et al., 2010). In 2015, Nickelsen et al. (2015) introduced a full dynamic iron module into the KMBM1 (named KMBM2), which put the number of biogeochemical parameters at 46 (more than doubled compared to the original).

Meaningful projections of ocean system response to climate change are only possible if key system processes and feedbacks are implemented (Anderson, 2005; Doney, 1999; Pomeroy, 2001; Denman, 2003). Hence, a model such as KMBM, which is widely used in studies of the present and paleo ocean (e.g., Meissner et al., 2005; Koeve and Kähler, 2010; Landolfi et al., 2013; Bralower et al., 2014; Kvale et al., 2018; Oschlies et al., 2019) and assessment of different geo-engineering measures under climate change (e.g., Reith et al., 2016; Mengis et al., 2019), has increased in complexity with time. However, increased complexity does not automatically guarantee a more realistic model, e.g., a model that better reproduces certain tracer distributions, surface fields, and fluxes.

The iron model is a good example of this. A fully dynamic iron cycle might be necessary to evaluate climate change because the future is predicted to become "dustier" (Hamilton et al., 2020) due to increased fossil fuel emissions and wildfire events. A dustier future suggests a dynamic dust flux may be important to resolve, as well as a dynamic response of ligands, which keep iron in solution. However, it is difficult for current iron models to reproduce observational data (Tagliabue et al., 2016). This might have structural reasons (relevant processes are missing) or it might be due to wrong model parameters. Increasing

complexity requires careful parameter validation with independent global data sets (Anderson, 2005). Yet as model complexity increases, our ability to hand-tune parameters is challenged and efforts in parameter calibration are sometimes underwhelming; for example, in KMBM2, the only tuning performed was against surface ocean nitrate, phosphate, and iron concentration (Nickelsen et al., 2015).

One of the hurdles facing thorough model validation with global data sets is the millennial time scales involved in whole ocean biogeochemical equilibration. However, in coarse-resolution ocean models (i.e. those with horizontal resolutions of  $> 1^{\circ}$  and between 15 and 30 vertical layers), model integration time for individual full ocean model spin-ups can be reduced to the order of days (vs. weeks, months) when using "offline" methods. These methods use prescribed, seasonally-cyclic but otherwise constant circulation, as well as prescribed atmospheric boundary conditions. An example is the TMM (Khati-wala, 2007), which provides an opportunity to assess a wide range of parameters against global climatological data.

As a first step, this chapter systematically examines the sensitivity of parameters in KMBM2 against global data sets of nitrate, phosphate, oxygen, and iron. The uncertainty of parameter values is assessed in order to identify suitable parameters from the model for future calibration, which will also explore effects on other metrics, such as biogeochemical fluxes and model response to transient forcing.

#### 2.2. Methods

#### 2.2.1. The University of Victoria Earth System Climate Model version 2.9

UVic 2.9 is an earth system model of intermediate complexity. It combines a simple onelayer atmospheric model, a terrestrial model, a sea-ice model, and a three-dimensional general circulation ocean model. All of the components have a  $3.6 \times 1.8$  horizontal grid resolution. The ocean model has 19 vertical layers of increasing width towards the deep. In this study, we use an Upwind-biased scheme (UW3) (Holland et al., 1998; Griffies et al., 2004) instead of the second-order Flux-Corrected Transport scheme (FCT) in Nickelsen et al. (2015) in order to obtain a comparable circulation between the "online" and "offline" models. The vertical diffusion coefficient is increased from 0.15 to 0.43  $cm^2 s^{-1}$ , in order to counteract the strength reduction of the meridional overturning circulation caused by switching advection schemes. The comparison between the "online" and "offline" models can be found in Kvale et al. (2017). The biogeochemical model used in this study is KMBM2 (Nickelsen et al., 2015).

Parameter	Description	Value	Units	
1.Phytoplanl	$xton\left(P_O, P_D\right)$			
a	Maximum growth rate constant at $0^{\circ}C$	0.6	$d^{-1}$	
$m_{P_O}$	<i>P</i> <sub>O</sub> mortality rate	0.03	$d^{-1}$	
$m_{P_D}$	<i>P<sub>D</sub></i> mortality rate	$\mu_P^*$	$d^{-1}$	
$\mu_P^*$	Microbial loop recycling rate	0.015	$d^{-1}$	
$c_D$	Diazotrophs' handicap	0.4		
light limitati	on			
k <sub>w</sub>	Light attenuation due to water	0.04	$m^{-1}$	
k <sub>c</sub>	Light attenuation through phytoplankton	0.047	$m^{-1} (mmol N m^{-1})^{-1}$	
k <sub>I</sub>	Light attenuation through sea ice & snow	5.0	$m^{-1}$	
PAR	Fraction of photosynthetically active radiation	0.43		
$\theta_{min}$	Maximum Chl:C ratio, extreme iron limitation	0.01	$gChl(gC)^{-1}$	
$\theta_{max}$	Maximum Chl:C ratio, abundant iron	0.04	$gChl(gC)^{-1}$	
$\alpha_{min}$	Minimum chlorophyll specific initial slope in PI-curve	18.4	$\mu g C (g C h l)^{-1} (W m^{-2})^{-1} s^{-1}$	
$\alpha_{max}$	Maximum chlorophyll specific initial slope in PI-curve	73.6	$\mu g C (g Chl)^{-1} (W m^{-2})^{-1} s^{-1}$	
Macro-nutrie	ent limitation			
K <sub>N</sub>	Half saturation constant for N uptake	0.7	$mmol N m^{-3}$	
$R_{P\cdot N}$	Molar P:N ratio	0.0625		
$K_P$	Half saturation constant for P uptake	$K_N R_{P \cdot N}$	$\mu mol m^{-3}$	
Iron limitatio	on l			
$K_{F}^{P}$ min	Minimum $K_{F_{r}}^{P}$	0.04	$\mu mol Fe m^{-3}$	
$K_{F}^{P} max$	Maximum $K_{E}^{P}$	0.4	$\mu mol Fe m^{-3}$	
Pmax	$P_{\Omega}$ biomass above which $K_{F_{\Omega}}^{P}$ increases	0.15	$mmol N m^{-3}$	
$K_{E}^{D}$	Half-saturation constant for Diaz Iron limitation	0.1	$\mu mol Fe m^{-3}$	
R <sub>Fe</sub> N	Molar Fe:N ratio for iron uptake	66.25	$\mu mol Fe (mol N)^{-1}$	
2 Zooplankt	$ron(\mathbf{Z})$	00.20		
2. 200prank	$\frac{1}{10000000000000000000000000000000000$	0.4	<u>d</u> -1	
8Z	Z growth officiancy	0.4	a	
$g_Z$	Z growin enciency	0.4		
Ŷ	Z assimilation enclency	0.7	$(mm \circ l N m^{-3})^{-2} d^{-1}$	
m <sub>Z</sub> V	z quadratic moltanty	0.00	$(mmot N m^{-3})$ a	
<b>Λ</b> <sub>Graze</sub>	nall-saturation constant for Z grazing	0.15	mmoi in m	
$\Psi_{P_O}$	Z preference for $P_O$	0.5		
$\Psi_{P_D}$	Z preference for $r_D$	0.1		
Ψz	Z preference for Other Z	0.5		
ΨDet	Z preference for <i>Det</i>	0.3		
3. Detritus (I	Det)			
$\mu_{Det}^0$	Det remineralization rate at $0^{\circ}C$	0.055		
$\omega_{Det}^0$	Sinking speed of <i>Det</i> at surface	14	$md^{-1}$	
$\omega_{Det}^{\iota}$	Depth dependent sinking speed slope	0.06	$d^{-1}$	
Carbonate				
$R_{C:N}$	Molar $C: N$ ratio	6.625		
$R_{CaCO_3:POC}$	CaCO3 over nonalgal POC production ratio	0.03		
D <sub>CaCO3</sub>	CaCO3 remineralisation e-folding depth	6500	m	
4.Iron ( <i>Fe</i> )				
L <sub>T</sub>	Total ligand concentration	1	$\mu$ mol lig m <sup>-3</sup>	
$K_{FeL}$	Fe-ligand stability constant	$10^{5.5}$	$(mmol lig m^{-3})^{-1}$	
$R_{Fe:P_{sed}}$	Molar Fe: P for sedimentary iron source	0.004		
$O_{2min}$	Minimum O2 concentration for aerobic oxidation	5	$mmol O_2 m^{-3}$	
$K_{Fe}^{org}$	Organic-matter dependent scavenging rate	0.45		
$K_{Fe}^{org}s$	Organic-matter particle scaling for scavenging	0.58		
$K_{E_{e}}^{col}$	Fe colloidal production and precipitation rate	0.005	$d^{-1}$	
5. Oxvgen (4	2			
Row	Molar $Q \cdot N$ ratio	10		
6 Tomporate	rea danandanay	10		
0. remperati	Soling footon h for terreneture days a first second	1.044		
υ	Scaling factor o for temperature dependent process	1.000	(°C)-1	
С	scaling factor c for temperature dependent process	1.0	( U) ·	

**Table 2.1.:** The biogeochemical parameter values for KMBM2, summarized from Nickelsen et al. (2015).

#### 2.2.1.1. Iron model

The iron cycle model in KMBM2 largely follows Parekh et al. (2004, 2005) and Galbraith et al. (2010). Iron limitation affects phytoplankton growth through lower light harvesting ability and lower nutrient-dependent growth rate. The global ligand concentration is constant, and dissolved iron that does not bind with ligands is subjected to organic scavenging and colloid formation. The iron cycle in this study includes different iron sources, such as aerosol deposition, sediment release, and hydrothermal iron flux. The aerosol deposition of iron is estimated from the monthly output of an atmospheric dust transport model (Luo et al., 2008). The release of iron from sediment is linked to particulate organic phosphorus that reaches the sediment with a fixed ratio (Nickelsen et al., 2015; Elrod et al., 2004). Hydrothermal iron flux data is compiled by Tagliabue et al. (2010). A more detailed description of the iron cycle in KMBM2 can be found in Nickelsen et al. (2015).

#### 2.2.2. Transport matrix method

The offline version of UVic 2.9 (Kvale et al., 2017) is used here. However, the extraction of TMs for UVic 2.9 is not straightforward, and a few modifications have been made to the standard fully-coupled (online) model. First, in previous studies (e.g., in Keller et al., 2012; Nickelsen et al., 2015), UVic 2.9 uses the non-linear FCT advection scheme, while TMM requires a linear advection scheme. Kvale et al. (2017) extracted the TMs with UW3. Secondly, the TMs combine both the horizontal advection and diffusion terms and use a slightly larger time step of 28800 s (versus 27000 s in the online model) in order to create a good balance between stability and accuracy. A comparison of the online and offline model performance is described in detail in Kvale et al. (2017).

#### 2.2.3. Experiment design

There are forty-six biogeochemical parameters in KMBM2 and they can be divided into groups based on their domain (e.g., phytoplankton, zooplankon, and detritus; See Tab. 2.1). For each group of parameters a meta-analysis of the uncertainty in observational or synthetic values is conducted (See section 2.3) to select a few parameters for sensitivity analysis. For the sensitivity analysis, an upper and a lower boundary of a parameter is defined based on its field measurements, synthetic estimation, or variations from its reference values (See Tab. 2.2). For each selected parameter, two model simulations are set up called "high-cast" and "low-cast" using its upper and lower boundaries, respectively, while keeping all other parameters at their reference values. Then, using Taylor plots, the sensitivity of the model misfit against observational data to each parameter is assessed. The reference values of parameters are from the work of Nickelsen et al. (2015) and the
Physiological traits and light harvesting of phytoplankton         a       0.3       0.75 $m_{P_0}$ 0.015       0.06 $\mu_P^p$ 0.0075       0.03 $k_c$ 0.001       0.06 $\alpha_{max}$ 36.8       147.2         Phytoplankton nutrient dependency       Kn       0.35       8.0 $R_{P:N}$ 0.031       0.094       K_Femax       0.04       2.28 $P_{max}$ 0.075       0.3       K_Fe       0.05       1.5 $R_{Fe}$ 0.05       1.5       R       R_Fe       0.05       1.5 $Z_{Fe}$ 0.2       0.8 $\gamma$ 0.5       0.72 $g_Z^{\ell}$ 0.2       0.8 $\gamma$ 0.5       0.72 $m_Z$ 0.03       0.12       KGraze       0.075       0.3         Detritus $\mu_{Det}^0$ 0.03       0.08 $\omega_{Det}^{0}$ 0.24       0.107 $R_{C:N}$ 3.31       13.25       13.25       13.25       13.25       14.5         Iron cycle $L_T$ 0.5       2 $K_{Fe}^{cong}$ 0.23       4.5	Parameter	Low-cast value	High-cast value
$a$ 0.3       0.75 $m_{P_0}$ 0.015       0.06 $\mu_P^*$ 0.0075       0.03 $k_c$ 0.001       0.06 $\alpha_{max}$ 36.8       147.2         Phytoplankton nutrient dependency $K_N$ 0.35       8.0 $R_{P:N}$ 0.031       0.094 $K_{Fe}^{P}$ $M_{P:N}$ 0.031       0.094 $K_{Fe}^P$ 0.05       1.5 $R_{Fe}$ 0.05       1.5 $R_{Fe}$ 0.05       1.5 $Z_{OOplankton}$ $Z_{OOplankton}$ $Z_{OOplankton}$ $g_Z^0$ 0.2       1.9 $g_Z^e$ 0.2       0.8 $\gamma$ 0.5       0.72 $m_Z$ 0.03       0.12 $K_{Graze}$ 0.075       0.3         Detritus $\mu_{Det}^0$ 7       28 $\omega_{Det}^{Org}$ 7       28 $\omega_{Det}^{Org}$ 0.23       4.5 $K_{Fe}^{Fe}$ 0.29       0.95 $K_{Fe}^{Cold}$ 0.0025       0.05 $R_{Fe}^{Fe}$ 0.002 <td< td=""><td>Physiologic</td><td>cal traits and light</td><td>harvesting of phytoplankton</td></td<>	Physiologic	cal traits and light	harvesting of phytoplankton
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	а	0.3	0.75
$\mu_p^*$ 0.00750.03 $k_c$ 0.0010.06 $\alpha_{max}$ 36.8147.2Phytoplankton nutrient dependency $K_N$ 0.358.0 $R_{P:N}$ 0.0310.094 $K_{Fe}^p$ max0.042.28 $P_{max}$ 0.0750.3 $K_{Fe}^p$ 0.051.5 $R_{Fe:N}$ 33.1132.5Zooplankton1.9 $g_2^{e}$ 0.21.9 $g_2^{e}$ 0.20.72 $m_Z$ 0.030.12 $K_{Graze}$ 0.0750.3Detritus $\mu_{Det}^0$ 0.030.08 $\omega_{Det}^{pet}$ 728 $\omega_{Det}^{org}$ 0.234.5 $K_{Fe}^{org}$ 0.234.5 $K_{Fe}^{org}$ 0.290.95 $K_{Fe}^{col}$ 0.00250.05 $R_{Fe:P_{sed}}^{eff}$ 0.0020.008 $O_{2min}$ 2.510Oxygen cycle $R_{ON}$ 8.012.0	$m_{P_O}$	0.015	0.06
$k_c$ 0.001       0.06 $\alpha_{max}$ 36.8       147.2         Phytoplankton nutrient dependency $K_N$ 0.35       8.0 $K_N$ 0.35       8.0 $R_{P:N}$ 0.031       0.094 $K_{Fe}^pmax$ 0.04       2.28 $P_{max}$ 0.075       0.3 $F_{Fe}^max$ 0.05       1.5 $R_{Fe:N}$ 33.1       132.5         Zooplankton $g_Z^p$ 0.2       1.9 $g_Z^e$ 0.2       0.8 $\gamma$ 0.5       0.72 $m_Z$ 0.03       0.12 $K_{Graze}$ 0.075       0.3       Detritus $\mu_{Det}^0$ 0.03       0.08 $\omega_{Det}^p$ 0.03       0.04       0.107 $R_{CN}$ 3.31       13.25         Iron cycle $L_T$ 0.5       2 $K_{Fe}^{org}$ 0.23       4.5 $K_{Fe}^{org}$ 0.23       4.5 $K_{Fe}^{org}$ 0.29       0.95 $K_{Fe}^{ord}$ 0.0025       0.05 $R_{Fe}^{ord}$ 0.002       0.008 $\omega_{Dati}$ 0.002       0.008 $O_{2min}^{ord}$ 0.5 $Q$	$\mu_P^*$	0.0075	0.03
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	$k_c$	0.001	0.06
Phytoplankton nutrient dependency $K_N$ 0.35         8.0 $R_{P:N}$ 0.031         0.094 $K_{Fe}^{Pmax}$ 0.04         2.28 $P_{max}$ 0.075         0.3 $K_{Fe}^{P}$ 0.05         1.5 $R_{Fe:N}$ 33.1         132.5           Zooplankton $g_2^{\sigma}$ 0.2         1.9 $g_Z^e$ 0.2         0.8 $\gamma$ 0.5         0.72 $m_Z$ 0.03         0.12 $K_{Graze}$ 0.075         0.3           Detritus $u_{Det}^0$ 0.03         0.08 $\omega_{Det}^0$ 7         28 $\omega_{Det}^0$ 0.024         0.107 $R_{C:N}$ 3.31         13.25         Item cycle $L_T$ 0.5         2 $K_{Fe}^{org}$ 0.23         4.5 $K_{Fe}^{org}$ 0.23         4.5 $K_{Fe}^{org}$ 0.23         4.5 $K_{Fe}^{ord}$ 0.002         0.008 $O_{2min}$ 2.5         10           Oxygen cycle $K_{Fe}^{org}$ 8.0         12.0         12.0  <	$\alpha_{max}$	36.8	147.2
$K_N$ 0.35       8.0 $R_{P:N}$ 0.031       0.094 $K_{Fe}^Pmax$ 0.04       2.28 $P_{max}$ 0.075       0.3 $K_{Fe}^D$ 0.05       1.5 $R_{Fe:N}$ 33.1       132.5         Zooplankton $g_Z^0$ 0.2       1.9 $g_Z^e$ 0.2       0.8 $\gamma$ 0.5       0.72 $m_Z$ 0.03       0.12 $K_{Graze}$ 0.075       0.3         Detritus       0.03       0.12 $K_{Graze}$ 0.075       0.3         Detritus       0.03       0.08 $\omega_{Det}^0$ 7       28 $\omega_{Det}^0$ 0.024       0.107 $R_{CN}$ 3.31       13.25         Iron cycle $L_T$ 0.5       2 $K_{Fe}^{cold}$ 0.025       0.05 $K_{Fe}^{cold}$ 0.0025       0.05 $R_{Fe}^{cold}$ 0.002       0.008 $O_2min$ 2.5       10       Oxygen cycle $R_{ON}$ 8.0       12.0	Phytoplank	ton nutrient depen	ndency
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	K <sub>N</sub>	0.35	8.0
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$R_{P:N}$	0.031	0.094
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$K_{Fe}^{P}max$	0.04	2.28
$\begin{array}{c ccccc} K_{Fe}^{D} & 0.05 & 1.5 \\ R_{Fe:N} & 33.1 & 132.5 \\ \hline \text{Zooplankton} & & \\ \hline g_{Z}^{0} & 0.2 & 1.9 \\ g_{Z}^{e} & 0.2 & 0.8 \\ \gamma & 0.5 & 0.72 \\ m_{Z} & 0.03 & 0.12 \\ K_{Graze} & 0.075 & 0.3 \\ \hline \text{Detritus} & & \\ \hline \mu_{Det}^{0} & 0.03 & 0.08 \\ \omega_{Det}^{0} & 7 & 28 \\ \omega_{Det}^{1} & 0.024 & 0.107 \\ R_{C:N} & 3.31 & 13.25 \\ \hline \text{Iron cycle} & & \\ \hline L_{T} & 0.5 & 2 \\ K_{Fe}^{org} & 0.23 & 4.5 \\ K_{Fe}^{org} & 0.29 & 0.95 \\ K_{Fe}^{col} & 0.0025 & 0.05 \\ R_{Fe:P_{sed}} & 0.002 & 0.008 \\ O_{2min} & 2.5 & 10 \\ \hline Oxygen cycle & \\ \hline R_{O:N} & 8.0 & 12.0 \\ \hline \end{array}$	P <sub>max</sub>	0.075	0.3
$R_{Fe:N}^{}$ 33.1       132.5         Zooplankton $g_2^0$ 0.2       1.9 $g_Z^e$ 0.2       0.8 $\gamma$ $\gamma$ 0.5       0.72 $m_Z$ 0.03       0.12 $K_{Graze}$ 0.075       0.3         Detritus $\mu_{Det}^0$ 0.03       0.08 $\omega_{Det}^0$ 7       28 $\omega_{Det}^0$ 0.024       0.107 $R_{C:N}$ 3.31       13.25         Iron cycle $L_T$ 0.5       2 $K_{Fe}^{org}$ 0.23       4.5 $K_{Fe}^{org}$ 0.29       0.95 $K_{Fe}^{col}$ 0.0025       0.05 $R_{Fe}:P_{sed}$ 0.002       0.008 $O_2min$ 2.5       10         Oxygen cycle $R_{O:N}$ 8.0       12.0	$K_{Fe}^D$	0.05	1.5
Zooplankton $g_D^0$ 0.2         1.9 $g_Z^e$ 0.2         0.8 $\gamma$ 0.5         0.72 $m_Z$ 0.03         0.12 $K_{Graze}$ 0.075         0.3           Detritus	$R_{Fe:N}$	33.1	132.5
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	Zooplankto	on	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$g_Z^0$	0.2	1.9
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$g_{Z}^{\overline{e}}$	0.2	0.8
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	γ	0.5	0.72
$\begin{array}{c cccc} K_{Graze} & 0.075 & 0.3 \\ \hline \text{Detritus} & & & \\ \hline \mu_{Det}^{0} & 0.03 & 0.08 \\ \varpi_{Det}^{0} & 7 & 28 \\ \varpi_{Det}^{1} & 0.024 & 0.107 \\ R_{C:N} & 3.31 & 13.25 \\ \hline \text{Iron cycle} & & \\ \hline L_{T} & 0.5 & 2 \\ K_{Fe}^{org} & 0.23 & 4.5 \\ K_{Fe}^{org} s & 0.29 & 0.95 \\ K_{Fe}^{col} & 0.0025 & 0.05 \\ R_{Fe:P_{sed}} & 0.002 & 0.008 \\ O_{2min} & 2.5 & 10 \\ \hline \text{Oxygen cycle} & \\ \hline R_{O:N} & 8.0 & 12.0 \\ \hline \end{array}$	$m_Z$	0.03	0.12
Detritus $\mu_{Det}^0$ 0.03         0.08 $\omega_{Det}^0$ 7         28 $\omega_{Det}^1$ 0.024         0.107 $R_{C:N}$ 3.31         13.25           Iron cycle $L_T$ 0.5         2 $K_{Fe}^{org}$ 0.23         4.5 $K_{Fe}^{org}$ 0.29         0.95 $K_{Fe}^{col}$ 0.0025         0.05 $R_{Fe:P_{sed}}$ 0.002         0.008 $O_{2min}$ 2.5         10         Oxygen cycle $R_{O:N}$ 8.0         12.0	K <sub>Graze</sub>	0.075	0.3
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	Detritus		
	$\mu_{Det}^0$	0.03	0.08
	$\omega_{Det}^0$	7	28
$R_{C:N}$ 3.31       13.25         Iron cycle $L_T$ 0.5       2 $K_{Fe}^{org}$ 0.23       4.5 $K_{Fe}^{org}$ 0.29       0.95 $K_{Fe}^{orl}$ 0.0025       0.05 $R_{Fe:P_{sed}}$ 0.002       0.008 $O_{2min}$ 2.5       10         Oxygen cycle $R_{O:N}$ 8.0       12.0	$\omega_{Det}^{i}$	0.024	0.107
Iron cycle         0.5         2 $L_T$ 0.5         2 $K_{Fe}^{org}$ 0.23         4.5 $K_{Fe}^{org}s$ 0.29         0.95 $K_{Fe}^{col}$ 0.0025         0.05 $R_{Fe:P_{sed}}$ 0.002         0.008 $O_{2min}$ 2.5         10           Oxygen cycle $R_{O:N}$ 8.0         12.0	$R_{C:N}$	3.31	13.25
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	Iron cycle		
$\begin{array}{ccccc} K_{Fe}^{org} & 0.23 & 4.5 \\ K_{Fe}^{org} s & 0.29 & 0.95 \\ K_{Fe}^{col} & 0.0025 & 0.05 \\ R_{Fe:P_{sed}} & 0.002 & 0.008 \\ O_{2}min & 2.5 & 10 \\ Oxygen cycle \\ \hline \hline R_{O:N} & 8.0 & 12.0 \\ \end{array}$	L <sub>T</sub>	0.5	2
$K_{Fe}^{oorg}s$ 0.29         0.95 $K_{Fe}^{col}$ 0.0025         0.05 $R_{Fe:P_{sed}}$ 0.002         0.008 $O_{2min}$ 2.5         10           Oxygen cycle $R_{O:N}$ 8.0         12.0	$K_{Fe}^{org}$	0.23	4.5
$K_{Fe}^{col}$ 0.0025         0.05 $R_{Fe:P_{sed}}$ 0.002         0.008 $O_{2min}$ 2.5         10           Oxygen cycle         70.000         12.0	$K_{Fe}^{org}s$	0.29	0.95
$\begin{array}{cccc} R_{Fe:P_{sed}} & 0.002 & 0.008 \\ O_2min & 2.5 & 10 \\ Oxygen cycle \\ \hline R_{O:N} & 8.0 & 12.0 \\ \end{array}$	KE	0.0025	0.05
$\begin{array}{c} O_2 min & 2.5 & 10 \\ Oxygen cycle \\ \hline R_{O:N} & 8.0 & 12.0 \end{array}$	R <sub>Fe:Pead</sub>	0.002	0.008
$\frac{\bar{Oxygen cycle}}{R_{O'N}} = 8.0 \qquad 12.0$	$O_2 min$	2.5	10
$R_{\Omega \cdot N}$ 8.0 12.0	Oxygen cy	cle	
	$\frac{1}{R_{O\cdot N}}$	8.0	12.0

**Table 2.2.:** List of parameters in the sensitivity study and their low and high-cast values. The units of parameters are identical as in Tab. 2.1. For details about parameter values and references, please see the Results section below.

simulation with all reference values is named as "ref".

#### 2.2.3.1. Simulation setup

Only one parameter in each model run is varied at a time, while the others use the reference values in Nickelsen et al. (2015). Each model is integrated for 5000 model years with pre-industrial forcing. All model simulations were completed on the NEC HPC-Linux-Cluster at Christian-Albrechts-University in Kiel, which is built with Intel Xeon Skylake computer nodes, with 32 cores per node. Each simulation used three nodes, which took approximately six hours for a single 5000-year spin-up. In total, there are 55 spin-up simulations (2 for each of the 27 parameters selected plus 1 'ref' run).

#### 2.2.3.2. The Taylor plot

In order to compare performance of the model experiments I applied normalized Taylor diagrams (Taylor, 2001). It uses a single point on a polar-style plot displaying three different components: the correlation coefficient (*R*) between the model-predicted field and the observational field, the centered Root-Mean-Square Error (*RMSE*) between the two fields, and the ratio of the standard deviation of model results ( $\sigma_M$ ) and observational data ( $\sigma_O$ ). This is possible as the relationship between *R*, *RMSE*,  $\sigma_M$ , and  $\sigma_O$  follows the cosine law:

$$RMSE^2 = \sigma_M^2 + \sigma_O^2 - 2\sigma_M \sigma_O R \tag{2.1}$$

 $\sigma_M$ ,  $\sigma_O$ , and *RMSE* can be represented as the sides of a triangle. The angle between  $\sigma_M$  and  $\sigma_O$  equals  $cos^{-1}R$ . In a normalized Taylor diagram, the radial distances from the origin to the model represent normalized standard deviation ( $\sigma_M/\sigma_O$ ), the azimuthal position denotes the correlation between the model result and observational data. The observation is represented at the coordination of (1,0) and the distances between the model and the observation represent normalized *RMSE* (where *RMSE'* = *RMSE*/ $\sigma_O$ ). Model that is close to observation indicate a higher degree of model fitness (and a lower misfit). Note that in this study, model runs are compared with three observational data sets (nitrate, phosphate, and oxygen), hence each model run in a normalized Taylor plot is represented by three points. Also, the RMSE in this chapter is not volume-weighted as in chapter 3.

#### 2.2.3.3. Observational data

The observational data used in this study are annual mean concentrations of nitrate ( $NO_3$ ), phosphate ( $PO_4$ ), and oxygen ( $O_2$ ) from the World Ocean Atlas 2013 (Garcia et al., 2013b,a). The model is also compared to a dissolved iron (dFe) concentration database, which comprises a collection of observations from both the GEOTRACES Intermediate Data Product 2017 (7694 points; Schlitzer et al., 2018) and prior observations (12371 points) compiled by Tagliabue et al. (2012). This database was first regridded to the UVic 2.9 grid (5917 points). Most of the dissolved iron data are less than 13  $\mu mol m^{-3}$ . Six observations with unusually high dissolved iron concentrations from locations with high hydrothermal activity (up to 205  $\mu mol m^{-3}$ ) from the data set were excluded, as *RMSE* is highly sensitive to outliers.

#### 2.3. Results and discussion

#### 2.3.1. Phytoplankton parameters

Phytoplankton physiology

In KMBM2, the phytoplankton in-situ growth rate depends on light availability, macronutrient limitation, iron limitation, temperature, and the growth rate constant at 0°*C* (*a*; see equations 9-12 in Nickelsen et al., 2015). As summarised in Le Quéré et al. (2005), *a* ranges from 0.04 (for nitrogen fixer) to 0.6 (for pico-phytoplankton)  $d^{-1}$  at 0°*C* for different phytoplankton functional types. However, since the phytoplankton in KMBM2 represents all types of phytoplankton other than nitrogen fixers, we conducted the sensitivity tests on *a* with 0.3 and 0.75  $d^{-1}$ , which are 0.5 and 1.5 times of its reference value in Nickelsen et al. (2015). Diazotrophs have a lower growth rate because the process of  $N_2$ -fixation requires extra energy (25% energetically more costly than direct nitrate uptake from solution, Knapp et al., 2012) which is represented by a diazotroph's handicap constant ( $c_D$ , smaller than one). The growth rate constant at 0°*C* for diazotrophs is  $c_D$ times *a* and we kept  $c_D$  constant in the study.

Ordinary phytoplankton mortality  $(m_{P_0})$  is reported as  $0.014 \pm 0.002 \ d^{-1}$  under Nlimitation and about 0.005 to 0.067  $d^{-1}$  under P-limitation in a diatom culture study by Brussaard et al. (1997). Le Quéré et al. (2005) averaged  $m_{P_0}$  to 0.05  $d^{-1}$  based on data reported by Brussaard et al. (1997). Schartau and Oschlies (2003) constrained  $m_{P_0}$  as  $0.040 \pm 0.013 \ d^{-1}$  using a 1D-ecosystem model with observational data from the North Atlantic. The sensitivity high (low)-cast value for  $m_{P_0}$  is 1.5 (0.5) times of its reference value (0.03  $d^{-1}$  in Nickelsen et al., 2015).

There is no explicitly-modeled bacteria in KMBM2 and the process of recycling organic matter produced by phytoplankton back to dissolved nutrients by the microbial community is modeled as a constant microbial loop recycling rate ( $\mu_p^*$ , a sink term for the biomass of the ordinary phytoplankton, see Equation 4 in Keller et al., 2012) that is additionally modified by temperature. Although bacterial growth, which is related to the availability of labile dissolved organic carbon (Hansell, 2013), has been measured in different regions of the open ocean (Ducklow, 1999), there are no direct observational measurements of the integrated recycling ability of the microbial community, which could be used guide the choice of parameter values for  $\mu_p^*$ . Hence,  $\mu_p^*$  is usually seen as a tuning parameter. Note that in Nickelsen et al. (2015),  $\mu_p^*$  is also the mortality rate for diazotrophs and assumes that the modeled diazotrophs have their own designated mortality rate, but with the same value as  $\mu_p^*$ . The sensitivity high (low)-cast value for  $\mu_p^*$  is 1.5 (0.5) times of its reference value (0.015  $d^{-1}$  in Nickelsen et al., 2015).

Sensitivity tests are conducted on a,  $m_{P_0}$ , and  $\mu_P^*$  (Fig. 2.1). Within the casting ranges,



**Figure 2.1.:** Sensitivity of phytoplankton physiological parameters for different model tracers, represented by squares ( $NO_3$ ), stars ( $PO_4$ ), triangles ( $O_2$ ), and diamonds (dFe). Plot (a) is for the growth rate constant at 0°C, plot (b) is for the ordinary phytoplankton mortality and plot (c) is for the microbial loop recycling rate. Parameter values are represented by the colors purple (low-cast), green (ref), and yellow (high-cast). The observational data are located at point (1,0) of the azimuthal coordination and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal RMSE and the interval between two neighboring line is 0.1.

 $m_{P_0}$  is not sensitive against the global distributions of  $NO_3$ ,  $PO_4$ ,  $O_2$ , and dFe. A lower value of  $\mu_p^*$  and a higher value of *a* can decrease global *RMSE* at the expense of decreased model deviation, which means that the model result has a smaller deviation than shown from the observational data.

#### Light limitation

In KMBM2, four physical parameters regulate how much radiation reaches the phytoplankton in water; the fraction of photosynthetically active radiation (*PAR*), light attenuation by water ( $k_w$ ), attenuation by sea ice and snow ( $k_I$ ), and attenuation by phytoplankton ( $k_c$ ). *PAR* is estimated to be around 43 to 52% of the total photosynthetically active radiation reaching the ground from satellite observation (Frouin and Pinker, 1995). In Nickelsen et al. (2015), *PAR* is set to be 43%.  $k_w$  is measured in the laboratory (Pope and Fry, 1997) and  $k_I$  is reported by Shirasawa et al. (2001). We kept *PAR*,  $k_w$ , and  $k_I$  constant in the study. The phytoplankton cells in the water can also attenuate light and  $k_c$  is usually described as attenuation per chlorophyll concentration, but in KMBM2,  $k_c$  is described as attenuation per biomass.  $k_c$  varies widely from 0.008 to 0.054  $m^{-1} (mmol N m^{-1})^{-1}$  and is highly model-dependent (Kvale and Meissner, 2017).  $k_w$  and  $k_c$  control the radiation attenuation due to water and chlorophyll independently, while  $k_I$  and *PAR* both impose multiplicative control on the overall photosynthetically active radiation amount (See Eq. 14 in Keller et al., 2012). We test the sensitivity of  $k_c$  applying a slightly larger range (0.001 to 0.06  $m^{-1} (mmol N m^{-1})^{-1}$ ) than the range in Kvale and Meissner (2017).

Phytoplankton light-harvesting capabilities and chlorophyll synthesis are influenced by iron limitation, as suggested by field and lab experiments (Hopkinson and Barbeau, 2007; Price, 2005). In KMBM2, both the chlorophyll-to-carbon ratio ( $\theta$ , varies between  $\theta_{min}$  and  $\theta_{max}$ ) and the initial slope of the PI curve ( $\alpha$ , varies between  $\alpha_{min}$  and  $\alpha_{max}$ ) scale with an iron limitation factor (*felim<sub>O</sub>* and *felim<sub>D</sub>* for the ordinary phytoplankton and diazotroph respectively), which scales between 0 (extreme iron limited) and 1 (iron abundant):

$$J_{OI} = \frac{J_O^{max} \alpha_O \theta_O I}{[(J_O^{max})^2 + (\alpha_O \theta_O I)^2]^{1/2}},$$
(2.2)

$$J_{DI} = \frac{J_D^{max} \alpha_D \theta_D I}{[(J_D^{max})^2 + (\alpha_D \theta_D I)^2]^{1/2}},$$
(2.3)

$$\theta_O = \theta_{min} + (\theta_{max} - \theta_{min}) felim_O, \qquad (2.4)$$

$$\theta_D = \theta_{min} + (\theta_{max} - \theta_{min}) felim_D, \qquad (2.5)$$

$$\alpha_O = \alpha_{min} + (\alpha_{max} - \alpha_{min}) felim_O, \qquad (2.6)$$

$$\alpha_D = \alpha_{min} + (\alpha_{max} - \alpha_{min}) felim_D, \qquad (2.7)$$

$$felim_O = \frac{Fe}{K_{Fe}^P + Fe},\tag{2.8}$$

$$felim_D = \frac{Fe}{K_{Fe}^D + Fe},\tag{2.9}$$

where  $J_{OI}$  and  $J_{DI}$  are the light dependent growth rate for the ordinary phytoplankton and diazotroph respectively,  $J_O^{max}$  and  $J_D^{max}$  are the maximum potential growth rate for the ordinary phytoplankton and diazotroph respectively.

This formulation enhances light harvesting ability of phytoplankton when iron is abundant and it intensifies the light limitation for phytoplankton that is under iron stress in order to represent co-limitation by iron and light for phytoplankton. The observational value of  $\theta$  is between 0.007 and 0.072  $gChl(gC)^{-1}$  and  $\alpha$  between 6.4 and 100  $\mu gC(gChl)^{-1}(Wm^{-2})^{-1}s^{-1}$  (both compiled by Geider et al., 1997). Later, Galbraith et al. (2010) decreased the variability of both parameters in the process of model tuning, suggesting a range of 0.01 (min) to 0.04 (max)  $gChl(gC)^{-1}$  and 18.4 (min) to 73.6  $\mu gC(gChl)^{-1}(Wm^{-2})^{-1}s^{-1}$ , respectively. From Equations 2.2 and 2.3, it is apparent that  $\theta$  and  $\alpha$  are mathematically equivalent in the model. We doubled the range of  $\alpha$  from Galbraith et al. (2010) and used it as the test boundary for  $\alpha_{max}$  and kept  $\theta$  constant.

Fig. 2.2 shows that higher  $k_c$  and lower  $\alpha_{max}$  can improve fit between model and observa-



**Figure 2.2.:** Sensitivity of phytoplankton light limitation parameters. The different model tracers are represented by squares ( $NO_3$ ), stars ( $PO_4$ ), triangles ( $O_2$ ), and diamonds (dFe). Plot (a) is for the light attenuation by phytoplankton and plot (b) is for the maximum initial slope of the PI curve. Parameter values are represented by the colors purple (low-cast), green (ref), and yellow (high-cast). The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal RMSE and the interval between two neighboring line is 0.1.

tional global distribution of  $NO_3$ ,  $PO_4$ , and  $O_2$ . However, iron global *RMSE* is insensitive towards both parameters.

#### Macro-nutrient limitation

The half-saturation constant  $K_N$  imposes restriction on the phytoplankton growth under N-limitation and also restriction on the phytoplankton growth under P-limitation through the Redfield ratio  $R_{P:N}$ , since in KMBM  $K_P = K_N R_{P:N}$ . The value of  $K_N$  varies from 0.1 to 8 *mmol N m*<sup>-3</sup> in a literature study of different phytoplankton functional types (Litchman et al., 2007). The observational value of  $K_P$  compiled by Geider et al. (1997) varies from 4 to 700  $\mu$ mol m<sup>-3</sup>. Geider and La Roche (2002) pointed out that  $R_{N:P}$  in phytoplankton can be pushed to extreme values of less than 5 or above 100 related to the ratio of fixed N to P in cultures. Even under nutrient-replete growth conditions,  $R_{N:P}$  may vary from 5 to 19.

The test range of  $K_N$  is combined with the half of the reference value in Nickelsen et al. (2015) as low-cast and its higher value in Litchman et al. (2007) as high-cast. The sensitivity high (low)-cast value for  $R_{N:P}$  is 1.5 (0.5) times of its reference value (0.015  $d^{-1}$  in Nickelsen et al., 2015). The sensitivity test result of  $K_N$  and  $R_{N:P}$  is in Fig. 2.3. A greater  $K_N$  can improve the model *RMSE* in the field of  $NO_3$ ,  $PO_4$ , and  $O_2$ , but not *dFe*. A change in the value of  $R_{N:P}$  can improve *RMSE* for one tracer while worsening it for others.

#### Iron limitation

From Equation 2.3, we can see that  $K_{Fe}^D$  describes the iron limitation of diazotrophs following a Monod function outlined in Aumont and Bopp (2006). This is not the same case



**Figure 2.3.:** Sensitivity of phytoplankton macro-nutrient limitation parameters. The different model tracers are represented by squares ( $NO_3$ ), stars ( $PO_4$ ), triangles ( $O_2$ ), and diamonds (dFe). Plot (a) is for the half-saturation constant of N and plot (b) is for the ratio of P and N in phytoplankton. Parameter values are represented by the colors purple (low-cast), green (ref), and yellow (high-cast). The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal RMSE and the interval between two neighboring line is 0.1.

for the ordinary phytoplankton in KMBM2:

$$P_1 = min(P_O, P_{max}), P_2 = max(0, P_O - P_{max}), K_{Fe}^P = \frac{K_{Fe}^P minP_1 + K_{Fe}^P maxP_2}{P_1 + P_2}, \quad (2.10)$$

where  $P_O$  is the concentration of ordinary phytoplankton.  $K_{Fe}^Pmin$ ,  $K_{Fe}^Pmax$  and  $P_{max}$  describe the iron limitation of ordinary phytoplankton with a flexible iron half-saturation constant ( $k_{Fe}^P$ ), reflecting the observation that larger cells have higher iron half-saturation constants than smaller ones (Timmermans et al., 2004). The increase of biomass for the ordinary phytoplankton in KMBM2 is assumed to be associated also with an increase in mean cell size just as in PISCES (Aumont and Bopp, 2006). The ordinary phytoplankton, when below a threshold biomass concentration of  $P_{max}$ , is considered to be small nanophytoplankton and thus has a smaller iron half-saturation constant  $K_{Fe}^Pmin$ . When the ordinary phytoplankton continues growing and its biomass surpasses the threshold  $P_{max}$ , its  $k_{Fe}^P$  starts increasing towards  $K_{Fe}^Pmax$  (See Equation 2.10).

The  $k_{Fe}^{P}$  for large open-ocean diatoms from the Southern Ocean is reported ranging from 0.19 to 1.14  $\mu mol Fe m^{-3}$  in an incubation study (Timmermans et al., 2004). Price et al. (1994) reported an effective phytoplankton community  $K_{Fe}^{P}$  between 0.035 and 0.22  $\mu mol Fe m^{-3}$  in the equatorial Pacific Ocean. For  $K_{Fe}^{P}max$ , we use the reference value of  $K_{Fe}^{P}min$  as low-cast and the double of its value in observation as high-cast. The sensitivity high (low)-cast value for  $P_{max}$  is 1.5 (0.5) times of its reference value (0.015  $d^{-1}$  in Nickelsen et al., 2015).

Diazotrophs have higher iron demand because iron is needed in enzymes related to nitrogen fixation. Bucciarelli et al. (2013) determined a  $K_{Fe}^D$  of 13.9  $\pm$  3.3  $\mu$ mol Fe m<sup>-3</sup> for filamentous diazotrophs and Jacq et al. (2014) reported a  $K_{Fe}^D$  of 6.95  $\pm$  0.66  $\mu$ mol Fe m<sup>-3</sup>



**Figure 2.4.:** Sensitivity of phytoplankton iron limitation parameters. The shapes square, star, triangular and diamond denote  $NO_3$ ,  $PO_4$ ,  $O_2$ , and dFe respectively. The color purple, green, and yellow represents the low-cast, ref, and high-cast respectively. Plot (a) is for the maximum iron half-saturation constant of Fe for the ordinary phytoplankton, plot (b) is for the threshold biomass concentration for the ordinary phytoplankton, plot (c) is for the iron half-saturation constant of Fe for the ratio of Fe and N in phytoplankton. The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal RMSE and the interval between two neighboring line is 0.1.The observations are located at the point [1,1] and closer a model to this point, the better its performance.

for unicellular diazotrophs. Moore et al. (2008) estimated  $K_{Fe}^D$  to be 0.06  $\mu$ mol Fe m<sup>-3</sup> in tuning the Biogeochemical Elemental Cycling model (BEC, Moore et al., 2004). Given the disagreement of  $K_{Fe}^D$  values between models and observation, we chose the test range of  $K_{Fe}^D$  by varying its reference value in (Nickelsen et al., 2015), 0.5 times for low-cast and 1.5 times for high-cast.

 $R_{Fe:N}$  describes the molar ratio of iron-to-nitrogen uptake for plankton. When there is excess iron in the water (e.g. in coastal regions), phytoplankton take up more iron than what is needed for their maximum growth (Kustka et al., 2003a; Iwade et al., 2006; Chen et al., 2011; Shire and Kustka, 2015). This produces uncertainty in the stoichiometry. Twining and Baines (2013) compiled  $R_{Fe:N}$  from observations, giving a range from 287.5 to 325  $\mu$ mol Fe (mol N)<sup>-1</sup> for ordinary phytoplankton and from 312.5 to 1938  $\mu$ mol Fe (mol N)<sup>-1</sup> for Trichodesmium ( $R_{P:N} = 16$  is applied for conversion here). The value for  $R_{Fe:N}$  in Nickelsen et al. (2015) is 66.25  $\mu$ mol Fe (mol N)<sup>-1</sup> and we vary this value with a factor of 0.5 for low-cast and 1.5 for high-cast in the sensitivity study.

We chose  $K_{Fe}^{P}max$ ,  $K_{Fe}^{D}$ ,  $P_{max}$ , and  $R_{Fe:N}$  for the sensitivity study due to the relatively large disagreement between model parameter value and the observation (See Fig. 2.4). Only a change in the value of  $R_{Fe:N}$  produces a change in the *RMSE* of *dFe*. However, greater  $K_{Fe}^{P}max$ ,  $K_{Fe}^{D}$ ,  $R_{Fe:N}$ , or smaller  $P_{max}$ , has the potential of increasing model fitness against global distributions of  $NO_3$ ,  $PO_4$ , and  $O_2$ .

#### 2.3.2. Zooplankton parameters



**Figure 2.5.:** Sensitivity of zooplankton parameters. The shapes square, star, triangular and diamond denote  $NO_3$ ,  $PO_4$ ,  $O_2$ , and dFe respectively. The color purple, green, and yellow represents the low-cast, ref, and high-cast respectively. Plot (a) is for the maximum potential feeding capability of zooplankton at 0°C, plot (b) is for the zooplankton growth efficiency, plot (c) is for the zooplankton assimilation efficiency, plot (d) is the zooplankton mortality rate and plot(e) is for the half-saturation constant of the zooplankton feeding rate. The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal RMSE and the interval between two neighboring line is 0.1.The observations are located at the point [1,1] and closer a model to this point, the better its performance.

There are nine parameters in KMBM2 that control zooplankton dynamics (See Tab. 2.1).

 $g_Z^0$  describes the maximum potential feeding capability of zooplankton at 0°*C*. Calbet and Landry (2004) compiled a database from dilution experiments and showed that the  $g_Z^0$  for micro-zooplankton ranges from 0.06 to 1.9  $d^{-1}$ .

 $K_{Graze}$  is a half-saturation constant of the feeding rate, which gives the zooplankton feeding rate a Holling II functional response according to the prey density. In different biogeochemical models,  $K_{Graze}$  ranges from 0.15 to 1 *mmol N m*<sup>-3</sup> (Scott et al., 2011). We vary the reference value of  $K_{Graze}$  from Nickelsen et al. (2015) with a factor of 0.5 for low-cast and 1.5 for high-cast.

Zooplankton growth efficiency  $(g_Z^e)$  varies between 0.3 and 0.43 and zooplankton assimilation efficiency  $(\gamma)$  varies between 0.5 and 0.72 in culture experiments conducted on six different species of micro-flagellates (Fenchel, 1982). Pelegrí et al. (1999) reported a  $g_Z^e$  of 0.22 and  $\gamma$  of 0.55 for nano-flagellates. We vary the reference value of  $g_Z^e$  from Nickelsen et al. (2015) with a factor of 0.5 for low-cast and 1.5 for high-cast.

Le Quéré et al. (2005) compiled a zooplankton mortality rate ( $m_Z$ ) of 0.058  $d^{-1}$  for mesozooplankton based on the data from Brussaard et al. (1997) and Hirst and Kiørboe (2002).  $m_Z$  ranges from 0.02 to 0.05  $d^{-1}$  in some model studies (summarized by Scott et al., 2011). KMBM2 uses a quadratic mortality of 0.06 ( $mmol N m^{-3}$ )<sup>-2</sup>  $d^{-1}$  in Nickelsen et al. (2015) and we vary this reference value of  $m_Z$  with a factor of 0.5 for low-cast and 1.5 for high-cast.

Copepods have a very diverse diet that is strongly dependent on what is available in the water (Kleppel, 1993). In KMBM2, the zooplankton grazing preference for ordinary phytoplankton ( $\psi_{P_0}$ ), for other zooplankton ( $\psi_Z$ ), and for detritus ( $\psi_{Det}$ ) are set to be equal in KMBM2. Only the zooplankton grazing preference for diazotrophs ( $\psi_{P_D}$ ) is set to be lower (see Tab. 2.1). This is because diazotrophs are considered "poor food" due their toxicity and unpalatability (Hawser et al., 1992). Also, there are records of large diazotroph blooms ending up ungrazed and decomposing (Sellner, 1997). However, some pelagic harpacticoid copepods have high ingestion rates on trichodesmium (Conroy et al., 2017). We kept all grazing preference parameters constant in the study.

Of the five zooplankton-related parameters tested, two ( $g_Z^0$  and  $K_{Graze}$ ) show potential for improving model fitness against global distributions of  $NO_3$ ,  $PO_4$ , and  $O_2$  (fig. 2.5).  $m_Z$  shows no sensitivity and the change in  $\gamma$  and  $g_Z^e$  are relatively small compared to  $g_Z^0$ .



**Figure 2.6.:** Sensitivity of detritus parameters. The shapes square, star, triangular and diamond denote  $NO_3$ ,  $PO_4$ ,  $O_2$ , and dFe respectively. The color purple, green, and yellow represents the low-cast, ref, and high-cast respectively. Plot (a) is for the detritus remineralization rate at 0°*C*, plot (b) is for the sinking speed of detritus at surface, plot (c) is for the depth dependent sinking speed slope and plot (d) is the carbon-to-nitrogen ratio for organic matter. The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal *RMSE* and the interval between two neighboring line is 0.1.

#### 2.3.3. Detritus parameters

In KMBM2, sources of detritus include the mortality of phytoplankton, the mortality of zooplankton, excretion of zooplankton, and sloppy feeding by zooplankton. The detritus sinks through the water column at an increasing speed with depth while it remineralizes into the dissolved nutrient pools. This generally mimics the Martin et al. (1987) flux attenuation behavior. For the detritus remineralization rate at  $0^{\circ}C$  ( $\mu_{Det}^{0}$ ), model synthetic values vary between 0.05 and 0.10  $d^{-1}$  (Fasham, 1995; Doney et al., 1996). Grossart and Ploug (2001) measured the degradation of aggregates formed by fresh phytoplankton detritus and reported a remineralization rate between 0.09 and 0.23  $d^{-1}$  at  $16^{\circ}C$ ; thus a range between 0.03 and 0.08  $d^{-1}$  at  $0^{\circ}C$  is used here.

 $\omega_{Det}^0$  and  $\omega_{Det}^i$  describe the detritus sinking speed as a linear function of depth (z),  $\omega_{Det} = \omega_{Det}^0 + \omega_{Det}^i z$ , mimicking the pattern observed by Berelson (2001). Berelson (2001) derived the sinking speed from sediment trap data in the Pacific and the Arabian Sea, thus providing a range between 0.024 and 0.107  $d^{-1}$  for  $\omega_{Det}^i$ . Combining the data sets from Berelson (2001) and Jackson et al. (2015) provides a range between 0 and 173  $md^{-1}$  for  $\omega_{Det}^0$ . The sensitivity high (low)-cast value for  $\omega_{Det}^0$  is 2 (0.5) times of its reference value

#### $(14 \ md^{-1} \text{ in Nickelsen et al., } 2015).$

The molar carbon-to-nitrogen ratio for organic matter ( $R_{C:N}$ ) in observations ranges from 3.8 to 12.5 (Geider and La Roche, 2002; Martiny et al., 2013; Singh et al., 2015). We did not conduct a sensitivity study on carbonate-related parameters because they do not influence the distribution of nitrate, phosphate, and oxygen in KMBM2. The diagnostic ratio of the carbonate formation to detritus formation ( $R_{CaCO_3:POC}$ ) is estimated to range between 0.03 and 0.09 in order to reproduce the observed alkalinity vertical profiles (Sarmiento et al., 2002; Jin. et al., 2006). The carbonate remineralization e-folding depth ( $D_{CaCO_3}$ ) is 6500 *m* in Nickelsen et al. (2015) and this parameter was found to be correlated with  $R_{CaCO_3:POC}$  (Ridgwell et al., 2007).

Four detritus parameters in KMBM2 are tested and only  $R_{C:N}$  is insensitive against the global *RMSE* of *NO*<sub>3</sub>, *PO*<sub>4</sub>, *O*<sub>2</sub>, and *dFe* (Fig. 2.6). There is no surprise that  $\mu_{Det}^0$ ,  $\omega_{Det}^0$ , and  $\omega_{Det}^i$  are sensitive for the distribution of the nitrate, phosphate, and oxygen (detritus is the link between surface biological activity and the deep ocean nutrient distribution). By changing any of those three parameters, the depth where detritus releases nutrient will change and so will the nutrient distribution. Also, remineralization of detritus consumes oxygen and the depth where remineralization happens, and at which rate it happens impacts the oxygen concentration. When the concentration of oxygen is reduced sufficiently, it creates the environment for denitrification, which is a loss term for the ocean nitrate inventory.

#### 2.3.4. Iron cycle parameters

Dissolved iron is assumed to have a chemical equilibrium between free dissolved iron, free ligands, and organic iron complexes. The global ligand concentration  $(L_T)$  is estimated to be around 1  $\mu$ mollig m<sup>-3</sup> (Parekh et al., 2005). In a database of ligand measurements compiled by Völker and Tagliabue (2015),  $(L_T)$  ranges from 0.5 to 6  $\mu$ mollig m<sup>-3</sup>. We vary the reference value (1  $\mu$ mollig m<sup>-3</sup> in Nickelsen et al., 2015) of  $L_T$  with a factor of 0.5 for low-cast and 2 for high-cast. The Fe-ligand stability constant ( $K_{FeL}$ ) is from 1011 to 1012 (mmollig m<sup>-3</sup>)<sup>-1</sup> in a review paper by Gledhill (2012). We kept  $K_{FeL}$  constant in the study.

Only the free dissolved iron is subject to organic scavenging (Parekh et al., 2005; Galbraith et al., 2010; Honeyman et al., 1988) and colloid formation. Honeyman et al. (1988) reported an organic-matter dependent scavenging rate  $(K_{Fe}^{org})$  of 0.079 and an organic-matter particle scaling for scavenging  $(K_{Fe}^{org}s)$  of 0.58, respectively derived by measurements of uranium (U) and thorium (Th) decay series isotopes. Lerner et al. (2016) reported the  $K_{Fe}^{org}$  and  $K_{Fe}^{org}s$  to be 6.62 and 0.73, respectively estimated from particle observational data in the North Atlantic. We apply a test range of 0.23 to 4.5 for  $K_{Fe}^{org}$ , which is half of and ten times its value in Nickelsen et al. (2015), respectively. For  $K_{Fe}^{org}s$ , we set a



**Figure 2.7.:** Sensitivity of iron cycle parameters. The shapes square, star, triangular and diamond denote  $NO_3$ ,  $PO_4$ ,  $O_2$ , and dFe respectively. The color purple, green, and yellow represents the low-cast, ref, and high-cast respectively. Plot (a) is for the global ligand concentration, plot (b) is for the organic-matter dependent scavenging rate, plot (c) is for the organic-matter particle scaling for scavenging, plot (d) is the iron colloid production and precipitation rate, plot (e) is for the ratio between sedimentary iron release and organic phosphate, and plot (f) is for the minimum  $O_2$  concentration for aerobic oxidation. The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal *RMSE* and the interval between two neighboring line is 0.1.

range between 0.29 and 0.95, which is half of its lower value in Honeyman et al. (1988) and about 30% larger than its value in Lerner et al. (2016).

There is no observational reference from iron colloid production and precipitation rate  $(K_{Fe}^{col})$ , since most of the estimates did not distinguish between iron scavenged by organic, versus inorganic (colloid formation) processes. Here, we apply a range from 0.0025 to 0.05  $d^{-1}$ , which is a factor of 0.5 and 10 of its reference value Nickelsen et al. (0.005  $d^{-1}$  in 2015).

However, both organic scavenging and colloid formation processes stop when the water becomes anoxic (which associates with a lower pH value). In KMBM2, the parame-

ter identifying the anoxic water is the minimum  $O_2$  concentration for aerobic oxidation  $(O_{2min})$ .  $O_{2min}$  is 5 mmol  $O_2 m^{-3}$  in KMBM2 as a pragmatic choice based on observations by Hopkinson and Barbeau (2007). Nickelsen et al. (2015) tested a range of values on  $O_{2min}$  and reported a minor impact. Note that in the model, denitrification also occurs in anoxic water, where the oxygen concentration is below  $O_{2min}$ . We vary the reference value of  $O_{2min}$  (5 mmol  $O_2 m^{-3}$  in Nickelsen et al., 2015) by a factor of 0.5 for low-cast and 2 for high-caster in the sensitivity study.

Elrod et al. (2004) reported a close dependency between iron flux from the sediments with the oxidation of organic matter, based on data from the California coast. The ratio between sedimentary iron release and organic phosphate ( $R_{Fe:P_{sed}}$  is 0.072. In KMBM2, the reference value for  $R_{Fe:P_{sed}}$  is 0.004 in Nickelsen et al. (2015). We vary its reference value by a factor of 0.5 for low-cast and 2 for high-cast in the sensitivity study.

From Fig. 2.7, we can see that five parameters,  $L_T$ ,  $K_{Fe}^{org}$ ,  $K_{Fe}^{org}s$ ,  $K_{Fe}^{col}$ , and  $R_{Fe:P_{sed}}$  have potential to improve the model *RMSE* for  $NO_3$ ,  $PO_4$ , and  $O_2$ . Smaller  $L_T$ ,  $K_{Fe}^{org}s$ , and  $R_{Fe:P_{sed}}$  can increase the correlation yet decrease deviation for the model, while larger  $K_{Fe}^{org}$  and  $K_{Fe}^{col}$  have the same effect. It seems that the parameter tuning for iron cycle in Nickelsen et al. (2015) is already quite good. However, only one parameter,  $L_T$  has a strong impact on the *RMSE* of *dFe*. This suggests a limited understanding of the global iron cycle (key processes are still missing from the model, or the modelling of certain processes is inadequate) with the additional challenge of still very sparse observational data coverage for iron in the ocean (Tagliabue et al., 2016).

#### 2.3.5. Oxygen parameter

In KMBM2, oxygen is produced during phytoplankton photosynthesis and consumed by organic matter remineralization at a molar ratio between oxygen and nitrogen ( $R_{O:N}$ ).  $R_{O:N}$  ranges from 8.1 to 11.2 for remineralization, based on data obtained in the Northeast Atlantic Ocean (Körtzinger et al., 2001) (assuming the average composition of phytoplankton and the respiration of organic matter with a combination of 52% protein, 36% polysaccharide, and 12% lipid). We tested the range of  $R_{O:N}$  from 8 to 12 in the study.

Unlike most of the previous parameters (e.g.  $L_T$  in Fig. 2.7) that improve the *RMSE* of  $NO_3$  and  $PO_4$  by increasing correlation and decreasing deviation at the same time,  $R_{O:N}$  (See Fig. 2.8) decreases the *RMSE* of  $NO_3$  and  $PO_4$  by increasing both correlation in  $NO_3$  and  $PO_4$ , yet increasing deviation of  $NO_3$  and decreasing deviation of  $PO_4$ . However, for the *RMSE* of  $O_2$ ,  $R_{O:N}$  behaves similarly to the previous parameter (e.g.  $L_T$  in Fig. 2.7).



**Figure 2.8.:** Sensitivity of oxygen cycle parameters: the molar ratio between oxygen and nitrogen. The shapes square, star, triangular and diamond denote  $NO_3$ ,  $PO_4$ ,  $O_2$ , and dFe respectively. The color purple, green, and yellow represents the low-cast, ref, and high-cast respectively. The observational data are located at point (1,0) and model performance is indicated by proximity to this point. The red dashed circles denote contours of equal *RMSE* and the interval between two neighboring line is 0.1.

#### 2.3.6. Temperature parameters

A lot of biogeochemical processes are temperature sensitive, e.g., phytoplankton growth, respiration, organic matter remineralization. Take the formulation of maximum potential growth rate of the ordinary phytoplankton  $(J_O^{max})$  and diazotroph  $(J_D^{max})$  in KMBM2 for example:

$$J_O^{max} = ab^{cT} felim_O, J_D^{max} = ab^{cT} felim_D,$$
(2.11)

where T is the in situ temperature. Scaling factors b and c for temperature-dependent processes originated in Eppley (1972). Together with a, they describe the relationship between phytoplankton growth rate and temperature. However, these two parameters characterize all temperature-dependent biological and chemical processes in KMBM2. The value of b and c were optimized in a 1D model with local data sets in Oschlies and Schartau (2005) and are left unchanged in this study.

#### 2.4. Conclusions

Increasing numbers of parameters present increasing challenges for parameter tuning, as model complexity increases. It is important to systematically assess existing parameter uncertainties and sensitivities.

In this chapter, it is shown that changing single parameter values within a tested range results in relatively similar model fitness (against global distributions of  $NO_3$ ,  $PO_4$ ,  $O_2$ , and dFe) compared to reference values that have been hand-tuned to reproduce surface nutrient distributions. This provides a perspective on the practise of model hand-tuning against surface nutrient fields, which for practical reasons is typically done on individual parameters. Nevertheless, a majority of parameters could potentially be calibrated to fur-

ther improve the model RMSE against observations.

Another key finding from this study is that most parameters are not able to improve model performance with respect to reproducing the global dissolved iron distribution. This may be due to poor model formulation of certain key processes, e.g., Völker and Tagliabue (2015) point out that an explicit dynamic ligand cycle in a biogeochemical model can improve the performance of the iron model. This might also be due to missing processes of the iron cycle, e.g., Beghoura et al. (2019) theorize missing fluxes of iron from the continental shelf. This might be due to the inadequate temporal and spatial coverage of Fe measurements. Caution is therefore advised when using Fe concentration data in global biogeochemical calibration studies in their current state.

Through the literature and sensitivity study, we prepared ourselves for the parameter calibration study presented in the next chapter. Because the calibration framework (Kriest et al., 2017) is not tested for calibrating more than six parameters at the same time, it was important to choose parameters that are sensitive for global distributions of nitrate, phosphate, and oxygen, e.g.,  $a, g_Z^0$ . We avoided choosing mathematically identical parameters to calibrate at the same time, e.g.,  $\theta$  and  $\alpha$ . We also chose a wide range of parameters from different domain (e.g., phytoplankton parameters, zooplankton parameters, detritus parameters, oxygen parameters), e.g., a parameter set as a (phytoplankton physiology),  $K_{Fe}^P$  (phytoplankton nutrient limitation),  $g_Z^0$  (zooplankton)  $\omega_{Det}^i$  (detritus),  $K_{Fe}^{org}$  (iron cycle), and  $R_{O:N}$  (oxygen).

# 3. Hierarchy of calibrated global models reveals improved distributions and fluxes of biogeochemical tracers in models with explicit representation of iron

This chapter is based on the paper "Hierarchy of calibrated global models reveals improved distributions and fluxes of biogeochemical tracers in models with explicit representation of iron" published in the journal Environmental Research Letters. Citation: Wanxuan Yao, Karin F Kvale, Eric Achterberg, Wolfgang Koeve, Andreas Oschlies (2019) Environ. Res. Lett. 14 114009, doi:10.1088/1748-9326/ab4c52.

#### 3.1. Abstract

Iron is represented in biogeochemical ocean models by a variety of structurally different approaches employing generally poorly constrained empirical parameterizations. Increasing the structural complexity of iron modules also increases computational costs and introduces additional uncertainties, with as yet unclear benefits. In order to demonstrate the benefits of explicitly representing iron, we calibrate a hierarchy of iron modules and evaluate the remaining model-data misfit. The first module includes a complex iron cycle with major processes resolved explicitly, the second module applies iron limitation in primary production using prescribed monthly iron concentration fields, and the third module does not explicitly include iron effects at all. All three modules are embedded into the same circulation model. Models are calibrated against global data sets of  $NO_3$ ,  $PO_4$  and  $O_2$  applying a state-of-the-art multi-variable constraint parameter optimization. The model with fully resolved iron cycle has a marginal better (up to 4.8%) skill in representing global distributions of  $NO_3$ ,  $PO_4$  and  $O_2$  compared to models with implicit or absent parameterizations of iron. We also found a slow down of global surface nutrient cycling by about 30% and a shift of productivity from the tropics to temperate regions for the explicit iron module. The explicit iron model also reduces the otherwise overestimated volume of suboxic waters, yielding results closer to observations.

#### 3.2. Introduction

Historically, global biogeochemical models have generally been becoming more complex. There is some justification for this; for example by a reduced misfit in global models using multiple phytoplankton functional types versus single-phytoplankton models (Friedrichs et al., 2007). However, increasing structural complexity requires careful assessment, if

anything is to be learned from the modelling study (Anderson, 2005). Models are also becoming more complex with respect to biogeochemical processes. Similar model-data misfits remaining after parameter optimization are possible for both structurally simple and complex biogeochemical ocean models (Kriest, 2017), but resulting nutrient pathways, i.e. routes taken by nutrients between particulate and dissolved pools, can differ substantially (Friedrichs et al., 2007; Löptien and Dietze, 2017). Biogeochemical differences arising due to differences in nutrient pathways can increase with climate change forcing (e.g. Kvale et al., 2015; Laufkötter et al., 2016), further underscoring the need for careful assessment as part of model development. Differences in nutrient pathways may also be relevant when modeling higher trophic levels, including fish (Pauly and Christensen, 1995; Stock et al., 2017).

Iron modeling forms an excellent example of the challenges mentioned above. An insufficient supply of the trace metal iron limits primary production in over one-third of the surface ocean (e.g. Boyd et al., 2007; Moore et al., 2013; Boyd and Ellwood, 2010) and also controls di-nitrogen fixation in nitrate-limited ocean regions (Schlosser et al., 2014). Iron is therefore considered an important regulator of the strength of the soft tissue pump (Volk and Hoffert, 1985) and, via atmospheric  $CO_2$ , of the climate (Joos et al., 1991). Consequently, earth system models (including all CMIP5 models; Laufkötter et al., 2015) include some parameterization of the marine iron cycle. However, a comparison of 13 global ocean biogeochemistry models against a compilation of dissolved iron (dFe) measurements showed that all of the models had clear deficits in reproducing many aspects of the observed patterns (Tagliabue et al., 2016). The authors emphasized that iron scavenging parameters are particularly poorly constrained. Iron scavenging (the adsorption of dFeonto particle surfaces) is commonly hand-tuned to achieve a good model fit to observed dFe concentrations (resulting in global mean model water column concentrations of 0.58  $\pm 0.14 \ \mu mol \ m^{-3}$  across the range of models, Tagliabue et al., 2016). Tuning a model to dFe using the iron scavenging parameterization has unclear consequences for model behavior. Pasquier and Holzer (2017) calibrated a simple steady-state model against multiple data fields (dFe,  $PO_4$ ,  $Si(OH)_4$  and chlorophyll) and found that they could achieve a similar model misfit (differing by less than 1%) when applying a wide range of assumed external iron source strengths, however, resulting in differing parameter values of the iron module. The Pasquier and Holzer (2017) study advanced our understanding of the sensitivity of simulated nutrient distributions to specific parameterizations of the iron cycle. The impact of including iron models of different structural complexity on the cycling of nutrients in a seasonally varying model environment is as yet un-quantified.

Nickelsen et al. (2015) introduced a dynamic iron module into an earth system model of intermediate complexity and hand-tuned model parameters against surface macro-nutrient observations and sparse observations of iron concentrations. The module resolved major components of the marine iron cycle, e.g. iron sources including aerosol deposition, detrital remineralization and sedimentary release, and iron sinks including biological uptake, iron scavenging and colloid formation. The focus of this earlier effort was on exploring

the impact of an explicit iron cycle on model sensitivities to environmental change, while the hand-tuning ensured that overall model performance for present-day ocean state was not affected negatively by the addition of the iron module.

In this study we present calibrations of three variants of a global model of ocean biogeochemical cycles (dynamic iron cycle, Nickelsen et al. 2015; prescribed iron mask, Keller et al. 2012; without iron, disabled iron limitation in Keller et al. 2012). Model variants share the same physical circulation but differ in their representation of the micro-nutrient iron. We calibrate against oceanic observations of  $NO_3$ ,  $PO_4$  and  $O_2$ , using a recently developed framework (Kriest et al., 2017). Our aim was to assess whether, and to what extent, the incorporation of Fe-related processes in the model improves the model skill of simulating  $NO_3$ ,  $PO_4$  and  $O_2$  as well as global indicators of biogeochemical cycles (described below). While it is generally assumed that the inclusion of an explicit iron cycle improves the capability of marine biogeochemical models to simulate distributions and fluxes of biogeochemical tracers also other than iron, this has, to our knowledge, not yet been demonstrated in a quantitative manner.

#### 3.3. Materials and Methods

#### 3.3.1. Model description

Our ocean model is a coupling of the University of Victoria Earth System Climate Model (UVic ESCM, version 2.9; Weaver et al., 2001; Eby et al., 2013) with the Transport Matrix Method (TMM; Khatiwala, 2007) architecture (an "offline" physical circulation, Kvale et al. 2017, see appendix). For the biogeochemical ocean model component, we deploy different versions of the Kiel Marine Biogeochemistry Module (KMBM), which differ in their representations of iron. The first one utilizes a full dynamic iron cycle (FeDyn; Nickelsen et al., 2015), while the second one utilizes prescribed monthly dFe surface concentration fields (FeMask; Keller et al., 2012), which is output from the Biology Light Iron Nutrient and Gas model (BLING), coupled with MOM4p1 circulation (Galbraith et al., 2010), regridded to UVic2.9 ESCM grid by Keller et al. (2012), and the third one includes no representation of the iron cycle (NoFe). NoFe is created by turning off the iron control on primary production in FeMask. Detailed descriptions of FeDyn can be found in Nickelsen et al. (2015); FeMask and NoFe in Keller et al. (2012). In FeDyn, we also include hydrothermal vents as an additional source of dFe to the deep ocean. The hydrothermal iron flux data was compiled by Tagliabue et al. (2010).

#### 3.3.2. Calibration framework

The calibration framework used in this study is adapted from Kriest et al. (2017). It contains three components; TMM, a biogeochemical model (either FeDyn, FeMask, or NoFe) and the Covariance Matrix Adaption Evolution Strategy (CMA-ES) (see appendix, Kriest et al., 2017; Hansen, 2006). CMA-ES is an evolutionary algorithm mimicking natural selection to search for an "individual" (a set of parameter values) that minimize the datamodel misfit. It starts with a random set of individuals (the first generation) and evaluates the model-data misfit of each individual. Individuals with lowest misfit are more likely to be present in the next generation, but random mutations are also permitted, that explore the search space with a mutation rate and intensity computed from the convergence behavior of previous generations. In our study, each individual is represented by a set of parameters used for a 3000 model-year model spin-up and subsequent computation of the model-data misfit. The CMA-ES stops when a pre-defined upper limit of the number of generations is reached or when successive generations do not yield a further reduction in the model-data misfit. The individual with the lowest model-data misfit is then the calibrated parameter set. All model spin-ups are performed with the same transport matrices as used by Kvale et al. (2017, derived from online simulations of the UVic ESCM; see appendix).

#### 3.3.3. Misfit function

Following Kriest et al. (2017), we define the misfit  $(J_T)$  of the model as the sum of the volume weighted Root Mean Square Error (RMSE) between simulated and observed annual mean concentrations of the tracers  $NO_3$ ,  $PO_4$  and  $O_2$  (World Ocean Atlas 2013; Garcia et al., 2013a,b). Volume weighting of the RMSE puts relatively more weight on the ocean interior than the surface ocean model-data misfit. The misfit function  $J_T$  is calculated as:

$$J_T = \sum_{j=1}^M J_j = \sum_{j=1}^M \frac{1}{\overline{o_j}} \sqrt{\sum_{i=1}^N (m_{i,j} - o_{i,j})^2 \frac{V_i}{V_T}}, M = 3$$
(3.1)

(3.2)

where j = 1, ..., M = 3 denotes tracers  $NO_3$ ,  $PO_4$  and  $O_2$ ; and i = 1, ..., N denotes the location of each model grid cell, N = 87307 for the UVic ESCM;  $V_T$  is the total volume of the model ocean and  $V_i$  is the volume of the respective grid cell;  $o_j$  is the global average observed concentration of the respective tracer;  $m_{i,j}$  and  $o_{i,j}$  are the modeled and observed concentrations of tracer *j* at location *i*, respectively. Note that, in order to allow for a fair comparison between models with and without an iron cycle, dFe is not included in the misfit function.

Parameter	Description	Observation	Cali. bound	uncalibrated value			Unit
				NoFe	FeMask	FeDyn	
a	Maximum growth rate at $0^{\circ}C$	0.04 - 0.60 <sup>a</sup>	0.30 - 0.75	0.600	0.600	0.600	$d^{-1}$
$\mu_P^*$	Microbial loop recycling rate		0.001 - 0.035	0.015	0.015	0.015	$d^{-1}$
$K_{Fe}^P$	Ordinary phyto. half-satu. of Fe-limitation	0.035 - 1.14 <sup>bc</sup>	0.04 - 1.14	-	0.1	-	$\mu$ mol Fe m $^{-3}$
$K_{Fe}^{P}max$	Maximum $K_{Fe}^P$	0.035 - 1.14 <sup>bc</sup>	0.04 - 1.14	-	-	0.4	$\mu$ mol Fe m $^{-3}$
$K_{Fe}^{D}$	Diazotrophs half-satu. of Fe-limitation	6.29 - 17.2 <sup>de</sup>	0.05 - 1.50	-	0.1	0.1	$\mu mol Fe m^{-3}$
$g_Z^0$	Maximum grazing rate at $0^{\circ}C$	0.06 - 1.9 <sup>f</sup>	0.2 - 2.0	0.4	0.4	0.4	$d^{-1}$
$\omega_{Det}^{i}$	Sinking speed slope	0.024 - 0.107 <sup>gh</sup>	0.001 - 0.160	0.06	0.06	0.06	$d^{-1}$
$K_{Fe}^{\overline{org}}$	POM dependent scavenging	0.079 - 6.62 <sup>ij</sup>	0.225 - 6.750	-	-	0.45	
$R_{O:N}$	Molar O : N ratio	$8.1 - 11.2^k$	8.0 - 12.0	10	10	10	
<sup><i>a</i></sup> Le Quéré et al. (2005). <sup><i>b</i></sup> Timmermans et al. (2004). <sup><i>c</i></sup> Price et al. (1994). <sup><i>d</i></sup> Bucciarelli et al. (2013) unpublished data.							
<sup><i>e</i></sup> Jacq et al. (2014). <sup><i>f</i></sup> Calbet and Landry (2004). <sup><i>g</i></sup> Berelson (2001). <sup><i>h</i></sup> Jackson et al. (2015). <sup><i>i</i></sup> Honeyman et al. (1988).							
<sup>j</sup> Lerner et al. (2016). <sup>k</sup> Körtzinger et al. (2001).							

**Table 3.1.:** The parameters chosen for each calibration, their observational value and calibration boundary settings. The parameters chosen for each calibration are marked with  $\sqrt{}$ .

#### 3.4. Results and discussion

#### 3.4.1. Improvement of misfit and biogeochemical indicators after calibration

For the dynamic iron model (FeDyn) we find a marked improvement by 8.6% for the total misfit after calibration, compared to the hand-tuned version of FeDyn (FeDyn0). Individual misfits  $J_{NO_3}$  and  $J_{PO_4}$  are reduced by 20.7% and 9.7%, respectively (Tab. 3.2), compared to FeDyn0. This significant gain in the model's ability to reproduce the observational pattern of NO<sub>3</sub> and PO<sub>4</sub>, comes at the expense of an increasing misfit  $J_{O_2}$ (5.8%), i.e. the ability to reproduce the observed patterns of  $O_2$ , after calibration. In our calibration, the objective is to minimize the overall misfit, which produced a tradeoff between individual components of the misfit function, leading to an increase in  $J_{O_2}$ . In order to assess whether a better overall misfit against observations of biogeochemical tracer concentrations improves the representation of the marine biogeochemical fluxes, we diagnosed several global indicators, i.e. net primary production (NPP,  $GtCyr^{-1}$ ), export production (Export,  $GtCyr^{-1}$ ), flux of organic carbon at a depth of 2 km ( $F_{POC}2km$ ,  $GtCyr^{-1}$ ), denitrification (Denit,  $PgNyr^{-1}$ ), the volume fraction of the ocean with oxygen concentrations less than 5  $mmol m^{-3}$  (OMZ5, %) and the volume fraction of ocean with oxygen concentrations less than 50  $mmol m^{-3}$  (OMZ50, %), and compared them to observational or synthetic data. Three out of six indicators (Denit,  $F_{POC}2km$ , OMZ5) show improvements for FeDyn against the hand-tuned FeDyn0 by getting closer to the observational data range (figure 3.1(a)). The estimations of Export are similar. Although NPP is increased by 4.5% after calibration, both the calibrated and hand-tuned models fall within the range of observational data. While the calibration improves the simulation of the OMZ5 by decreasing its volume by 73.6%, a reduction by 42.4% of OMZ50 size worsens the description of the volume fraction of hypoxic waters where  $O_2$  is less than  $50 \, mmol \, m^{-3}$ .

**Table 3.2.:** A comparison of the misfit from optimized model spin-ups. Values for uncalibrated FeDyn (FeDyn0) are calculated diagnostically (by us) and have not been used in the tuning process of Nickelsen et al. (2015). The minimum value of each column is marked in bold.

Experiment	$J_{NO_3}$	$J_{PO_4}$	$J_{O_2}$	$J_T$
NoFe	0.1575	0.1305	0.1786	0.4667
FeMask	0.1537	0.1293	0.1803	0.4632
FeDyn	0.1457	0.1254	0.1730	0.4442
Hand-tuned				
FeDyn0	(0.1838)	(0.1389)	(0.1635)	(0.4862)
-				



**Figure 3.1.:** Model global annual fluxes and oxygen minimum zone size against their respective observational data. We plot six global indicators on lineal axis pointing to the pole and the units are  $GtCyr^{-1}$  (NPP),  $PgNyr^{-1}$  (Denit),  $GtCyr^{-1}$  ( $F_{POC}2km$ ),  $GtCyr^{-1}$  (Export), %(OMZ5), % (OMZ50). The dash-dotted (dotted) line denotes the lower (upper) boundary of the observational value. The solid lines present the model annual mean values. (a) shows the comparison between un-calibrated FeDyn (magenta) with calibrated (blue); (b) shows the comparison between three calibrated models, FeDyn, FeMask (green) and NoFe (red); the observational data are from Carr et al. (2006, NPP), Kriest and Oschlies (2015, Denit, compilation of rates from other studies referred to in their Table 2), Honjo et al. (2008,  $F_{POC}2km$  and Export), Lutz et al. (2007, Export), Dunne et al. (2007, Export) and Garcia et al. (2013b, OMZs). Here we use 2.5 (7.5) mmol m<sup>-3</sup> as the lower (upper) boundary value for OMZ50 and 40 (60) mmol m<sup>-3</sup> as the lower (upper) boundary value for OMZ50 conversion between different elements followed C : N : P = 106 : 16 : 1.

#### 3.4.2. The calibrated dynamic iron model has the best skill

When comparing the misfits from different calibrations, FeDyn with a full dynamic iron cycle has a misfit  $J_T$  4.8% smaller than NoFe, and 4.1% smaller than FeMask (Tab. 3.2). Also for the individual components ( $J_{NO_3}$ ,  $J_{PO_4}$ , and  $J_{O_2}$ ) FeDyn provides the smallest misfit against observations. The most significant difference is in  $J_{NO_3}$ , with FeDyn having a 7.5 (5.2)% smaller  $J_{NO_3}$  than NoFe (FeMask). This improvement may be related to the improvement in OMZ5, and associated Denit, in FeDyn, described below. However, even the smallest  $J_T$  (FeDyn) still amounts to about 15% of global mean tracer concentrations of each tracer, which is on par with the calibration result from Kriest et al. (2017, conducted using a somewhat simpler biogeochemistry coupled to a different circulation) and considerably larger than the 5% obtained for  $PO_4$  by Pasquier and Holzer (2017). This is not unexpected because their study utilized an inverse model coupled with a data-assimilated, steady circulation (Primeau et al., 2013), where  $PO_4$  had been used already as part of the constraint for obtaining a circulation that can optimally represent the tracer field. The misfit for FeMask and FeDyn are around 15.4% and 15.6% of global mean tracer concentrations, respectively. This indicates that the differences between our biogeochemical models are smaller than the differences between either model and observations. We believe that this relates to biases in our physical circulation model (see figure A.1), which would also be consistent with the considerably smaller  $PO_4$  misfit obtained by Pasquier and Holzer (2017) for their PO<sub>4</sub>-constrained circulation field. In our model, annually-averaged basin vertical profiles of macro-nutrients and  $O_2$  (figure A.1) display the circulation bias reported in Kvale et al. (2017); a deep ocean (>2000 m) that is generally too old (according to the radiocarbon), which produces an overestimation of  $NO_3$  and  $PO_4$ , and an under-estimation of  $O_2$  in the deep ocean basins. The Pacific  $O_2$  profile is affected by an over-estimate of deep vertical mixing in the Southern Ocean (not shown). Even with calibration, these biases cannot be completely overcome.

For the six biogeochemical indicators, half of them (Denit, OMZ5 and NPP) improved with respect to observations after the inclusion of iron modules, two ( $F_{POC}2km$  and OMZ50) became worse and Export was essentially unchanged (figure 3.1(*b*)). The indicators most sensitive to iron are NPP and OMZ5: experiment FeDyn resulted in a reduction of NPP by 31.4% and OMZ5 by 71.9% compared to NoFe. Between the two models with iron, FeDyn shows lower  $F_{POC}2km$  (3.6%), OMZ5 (50.6%) and OMZ50 (23.5%) than FeMask. The former two values reflect improvements compared to observations, while the latter does not. When considering all indicators equally, the simulation of biogeochemical indicators is improved when iron modules are included, and a fully dynamic iron cycle outperforms a model with an iron mask. Both misfits (Tab. 3.2) and biogeochemical indicators are improved when a full dynamic iron cycle is used.

An interesting feature of the calibrations is the different behavior of OMZ50 and OMZ5 upon the inclusion of an explicit representation of the iron cycle, namely an improved agreement of the OMZ5 volume with observations while the agreement deteriorates for

the OMZ50 volume. Both volumes decreased after the calibration of biogeochemical parameters (improving the OMZ5 bias but exacerbating the underestimation of the OMZ50 volume with respect to NoFe). The OMZ volume is determined by a combination of physical and biogeochemical processes (Kriest et al., 2012; Kriest and Oschlies, 2015). The circulation is held constant across all model experiments, with only the biogeochemical parameters allowed to vary. The calibrations utilize global misfit weighted by ocean volume, which gives little consideration to the OMZ volume that comprises less than 4% of the global ocean. However, OMZ5 waters have a special role as featuring the only permanent sink of  $NO_3$  (via denitrification) in our model. The OMZ5 volume therefore strongly impacts the global nitrogen inventory and associated nitrate distributions (Landolfi et al., 2013; Kvale et al., 2019). The calibration against observed  $O_2$  and  $NO_3$  distributions emphasized the improvement of  $J_{NO_3}$ , via improved OMZ5 volume. In contrast, there is no such control via simulated  $NO_3$  distributions on the OMZ50 volume.

#### 3.4.3. Calibrated parameters and uncertainties

The calibrated parameter values and their range within a 1% increase in misfit around the minimum of the misfit function  $J_T$  for the respective calibrations define a measure of uncertainty of the calibrated parameters and are provided in Tab. 3.3. Also shown are the resulting uncertainties in NPP and Export. All parameters and fluxes are constrained within a relatively small range. The more tightly constrained parameters are the increase of sinking speed with depth and the molar O: N ratio. Their uncertainty ranges amount to 9.7-13.8% and 2.9-7.0% in the different calibrations, with calibration results for NoFe having the largest uncertainty. This is due to the higher misfit of NoFe, which, via our 1% misfit criterion increases the absolute misfit difference and therefore may lead to a larger range of parameter values consistent with a 1% misfit increase. The less constrained parameters are the microbial loop recycling rate, the diazotrophic half-saturation of iron and the iron scavenging rate. However, those parameters also have a relatively large prior parameter range used as input in the calibration. Interestingly, estimates of the tracer fluxes, such as NPP and Export are relatively well constrained, despite large uncertainties in individual parameters (bottom rows of Tab. 3.3). For each model, the uncertainty range in NPP is diagnosed from the ensemble of simulations and parameter combinations with a misfit function value less than 1% larger than the minimum of the cost function. The thus estimated uncertainties amount to 7.9-14.9% for NPP and 7.0-9.5% for Export (bottom rows of Tab. 3.3).

An encouraging result of the calibration is that NPP and Export, which are not explicitly included in the misfit function, turn out to be well constrained and in good agreement with independent observational evidence. Some optimal parameters appear to be portable across model configurations (e.g., see the close agreement between  $\omega_{Det}^i$  values). However, other parameters are model dependent (e.g., calibrated  $g_Z^0$  values vary up to 126% between models).

**Table 3.3.:** Calibrated parameter value and their resulting fluxes ranges with 1% of misfit. Here we show the range of parameter values and simulated NPP and Export diagnosed from all simulations that yield not more than a 1% increase of the misfit function with respect to its global minimum. The parameter values that are kept fixed during the calibration are put in parentheses. The unit of NPP and Export is  $Gt Cyr^{-1}$ . See Tab. 3.1 for the units of the individual parameters.

	NoFe		FeMask		FeDyn	
Parameter	value	range*	value	range*	value	range*
a	0.698	0.650 - 0.780	(0.060)		(0.060)	
$\mu_P^*$	0.0157	0.0142 - 0.0212	0.0012	0.0000 - 0.0020	(0.015)	
$K_{Fe}^P$	-		0.050	0.045 - 0.071	-	
$K_{Fe}^{P}max$	-		-		0.585	0.520 - 0. 598
$K_{Fe}^{D}$	-		0.406	0.387 - 0.494	0.377	0.326 - 0.479
$g_Z^0$	1.282	1.223 - 1.474	0.668	0.551 - 0.691	0.567	0.526 - 0.685
$\tilde{\omega_{Det}^i}$	0.065	0.063 - 0.072	0.062	0.058 - 0.064	0.060	0.059 - 0.065
$K_{Fe}^{org}$	-		-		0.427	0.208 - 0.510
$R_{O:N}$	10.44	10.18 - 10.91	9.54	9.34 - 9.62	10.502	10.27 - 10.58
Fluxes						
NPP	75.9	70.1 - 81.44	49.1	44.0 - 49.9	52.0	50.4 - 54.5
Export	6.3	5.8 - 6.4	6.8	6.4 - 7.0	7.1	6.7 - 7.2

\* Here is obtained by sampling model runs within 1% range of smallest misfit for the respected calibration

#### 3.4.4. The impact of including an iron module on surface ocean nutrient pathways

The different models reveal different surface ocean biogeochemical pathways (figure 3.2). Although the use of volume-weighted model-data misfits in the model calibration emphasizes fitness in the deep ocean, the biogeochemical parameters selected for the calibration reflect mostly upper-ocean processes. The differences in NPP are the most prominent, with NPP being much lower in FeDyn and FeMask (52.0 and 49.1  $GtCyr^{-1}$ ) compared to NoFe (75.9  $GtCyr^{-1}$ ). However, the corresponding total nutrient recycling (sum of the microbial loop recycling, zooplankton excretion and remineralization) also slows, which leads to relatively similar export production rates (6.3-7.1  $GtCyr^{-1}$ ) out of the euphotic zone (130 m) for all three calibrated models. The similarity of Export estimations between calibrations is likely linked to the calibration setup, which uses the same circulation to achieve the same observational distribution of tracer concentrations and, therefore, is likely to produce similar values of export production. Kriest (2017) calibrated two no-iron prognostic models with identical physical circulation against observational data, and found similar values of export production of between 6.0-7.4  $GtCyr^{-1}$ . In our calibrations, the inclusion of iron seems to slightly increase Export, nevertheless our global number (7.1  $GtCyr^{-1}$  in FeDyn) is still lower than the steady state model calibration from Pasquier and Holzer (2017, 9.5-11.0  $GtCyr^{-1}$ ) using a different circulation model. Our 1% misfit range of NPP and Export of respective calibrations can be found in Tab. 3.3.

The three calibrated model configurations result in similar Export through different sur-



**Figure 3.2.:** A NPZD plot comparing the surface nutrient pathway in three calibrated models. NoFe is red, FeMask is green and FeDyn is blue. The boxes represent nutrient (N), phytoplankton (P), zooplankton (Z) and detritus (D) and their inventory in the euphotic zone for each calibration are marked with numbers (unit PgN) in the box with respective color. The thicknesses of the arrows is approximately proportional to the associated flux and all fluxes are in unit of  $PgNyr^{-1}$ ).

face nutrient pathways. Zooplankton in our model grazes not only on phytoplankton but also on detritus and itself. The net nutrient flux from zooplankton to detritus represents sloppy feeding plus zooplankton mortality minus grazing on detritus. Increasing the zooplankton grazing rate increases the flux in both directions, with the net effect reducing the percentage of nutrient that goes directly from zooplankton to detritus, e.g., NoFe has the highest grazing rate and only 20.1% of NPP goes to detritus through zooplankton, comparing to 25.3% in FeMask and 27.9% in FeDyn. A higher grazing rate also reduces the standing stock of phytoplankton and reduces phytoplankton loss via mortality. Hence, despite the differences in NPP, detritus production rates are similar in the three calibrations (3.3  $PgNyr^{-1}$  in NoFe, 3.2  $PgNyr^{-1}$  in FeMask and 3.2  $PgNyr^{-1}$  in FeDyn). In all our calibrations, the remineralization rate at  $0^{\circ}C$  and its dependence on temperature is identical. The higher remineralization flux in NoFe is due to the high productivity in the tropical (warmer) region in that configuration (figure 3.3(a) and (d)). The higher remineralization of detritus keeps nutrients in the surface ocean. In combination with the higher phytoplankton growth rate and higher grazing rate this accelerates the surface ocean nutrient cycling (47.3% faster when comparing NoFe to FeDyn).

Differences in surface ocean nutrient cycling and pathways such as identified here may have substantial implications when projecting future changes to ecosystem services, e.g., sustainability and resilience of fisheries, since catches are closely related to NPP, trophic energy pathways and transfer efficiencies (Stock et al., 2017).

#### 3.4.5. Geographical shifts in NPP, Export and surface NO<sub>3</sub>

The geographical distributions of simulated marine biogeochemical fluxes also differs across calibrations. The NPP reductions per area are 150  $gCm^{-2}yr^{-1}$  for the Pacific, 30  $gCm^{-2}yr^{-1}$  for the Atlantic, and 60  $gCm^{-2}yr^{-1}$  for the Indian Ocean, when comparing FeDyn to NoFe. The Pacific Ocean is most impacted by iron limitation, since it hosts three major HNLC regions, the subpolar North Pacific, tropical Pacific, and part of the Southern Ocean (Pitchford, 1999; Cullen, 1995). However, the reductions of NPP in the Southern Ocean and North Pacific are small compared to the tropical Pacific (see figure (3.3(a)). The large reduction of NPP in the tropical Pacific in FeDyn is due to the reduction of nutrient recycling as compared to the NoFe model (described in section 3.4.4), where a higher phytoplankton growth rate and higher grazing rate accelerate nutrient recycling in the warm tropical waters in NoFe. This recycling is reduced in FeDyn via iron limitation on primary production, as well as this model having generally lower biological rates. Reduction of NPP in the tropical Pacific, caused by the inclusion of a dynamic iron module and regional iron limitation, results in higher annual mean euphotic zone NO<sub>3</sub> concentrations (by up to 8 mmol  $m^{-3}$ ) in the east Pacific upwelling region (see figure 3.3(c)). This residual  $NO_3$ , which has not been exported into the Pacific OMZ, can be utilized further downstream of the equatorial upwelling system, where iron is less limiting (e.g. the subtropical regions). Meanwhile, higher global Export in FeDyn (see figure 3.3(b))



**Figure 3.3.:** Difference of annual mean NPP, Export and euphotic zone  $NO_3$  concentration between calibrations. Note that after iron limitation is introduced, not all regions reveal a reduced NPP (e.g., temperate region in the Southern Ocean, North Pacific and East Equatorial Pacific). This is due to the strong grazing control on phytoplankton in those regions in NoFe.

produces lower annual-mean euphotic zone  $NO_3$  concentrations (by up to 14  $mmolm^{-3}$ ) in the rest of the surface ocean (e.g., equatorial Atlantic, see figure 3.3(c)).

Bopp et al. (2013) report that the different CMIP5 models show little agreement in simulated trends in primary production in eastern equatorial regions despite all models accounting for iron limitation. In our calibrated iron models (FeDyn and FeMask), we also see different changes in NPP and  $NO_3$  concentrations in this region (figure 3.3), which suggests that the available observations of  $NO_3$ ,  $PO_4$  and  $O_2$  together with a volumeweighted cost function might not be sufficient to constrain the model dynamics in this region.

#### 3.5. Conclusions and future work

This study describes a range of advantages of explicitly including iron in a global biogeochemical ocean model by comparing calibrated models that employ a hierarchy of different iron implementations. The models with explicit consideration of iron perform better compared to the model without iron in two respects: (a) improved representation of macro-nutrient and oxygen distributions with around a 4.5% reduction of total model misfit, and (b) generally improved estimations of global indicators of biogeochemical cycles. The inclusion of a calibrated iron module produces more realistic global NPP, which is around 33% smaller than in the model without iron. It also produces a more realistic suboxic volume ( $O_2$  is less than 5 mmol m<sup>-3</sup>), which is reduced by 72.0 (43.3) % in the dynamic iron model (iron mask model) compared to the no-iron model. A more realistic suboxic volume benefits the global nitrate distribution by constraining denitrification rates and is achieved in the iron models via a combination of changes in both tropical export production and oxygen demand for organic particle remineralization.

We therefore conclude that it is worthwhile to include a dynamic iron cycle in biogeochemical models for research questions relevant to distributions of macro-nutrients, oxygen, NPP and very low oxygen thresholds. However, modeling of water masses with somewhat higher oxygen concentrations (regions with  $O_2$  less than 50 mmol m<sup>-3</sup>), remains problematic. Furthermore, for each individual model, calibration improved the model performance against chosen metrics compared to the hand-tuned model, regardless of the structural model complexity. In our case, this improvement was generally larger than the differences between the calibrated models differing in the representation of the iron cycle. This is an important point, because not only does inclusion of iron improve the model performance, but also does the calibration itself offer substantial model improvements that cannot generally be achieved with hand-tuning. Systematic model calibration also reveals, which parameters are portable across biogeochemical model structures (in our instante, the rate of increase of the sinking speed with depth) and which are not (e.g., the grazing rate).

We also found that the iron scavenging parameter in our calibration is less well constrained then other parameters. This may be improved when we use direct iron measurements as an extra observational constraint. However, in order to better reproduce observed patterns of dissolved iron, we may need a more sophisticated ligands parameterization, as indicated by Völker and Tagliabue (2015). Ligands can keep dissolved iron in solution instead of being scavenged (Gledhill, 2012), which can directly impact the iron scavenging parameter estimation. As Buck et al. (2015, 2017) point out, ligands have different iron-binding strengths and life spans; Pham and Ito (2018, 2019) suggest ligands may be responsible for subsurface maxima of dissolved iron. Our model assumes a constant ligand concentration, which probably underestimates the control of ligands on our model biogeochemistry. It would be interesting to calibrate a more complex representation of ligands in the model in the future and to investigate if this can further improve the model performance.

The surface nutrient pathways differ across calibrated models, particularly with respect to total NPP. These differences arise from both the differences in model structure and the calibrated parameter values. Whether different nutrient pathways result in significantly different model responses to transient forcing (i.e. a changing climate) remains an open question. The fact that different optimal nutrient pathways arise in models of different complexity and model structure despite an identical calibration objective, demonstrates that surface processes such as microbial loop recycling cannot straightforwardly be constrained by seasonally cycling patterns of biogeochemical tracer concentrations. Lastly, the high sensitivity of simulated tropical Pacific NPP and Export to model structure emphasizes the importance of iron limitation in this region.

After the benefit of including an explicit representation of the iron cycle has been demonstrated in the current study, future work will investigate the effect of including direct iron measurements in the calibration process of such models. The impact of calibration will also be studied with respect to the simulated biogeochemical responses under a changing climate.

#### Acknowledgements

This work is supported by the Helmholtz Research School for Ocean System Science and Technology (HOSST, Germany) and its sister school, Transatlantic Ocean System Science and Technology (TOSST, Canada). We acknowledge discussions within the Biogeochemical Modelling and Marine Chemistry research units at GEOMAR, Dalhousie University, and University of Liverpool. We in particular thank Iris Kriest (GEOMAR), Alessandro Tagliabue (Liverpool) and Allison Chuan (Dalhousie) for continuous support. We also acknowledge the computing resources provided by the North German Supercomputing Alliance (HLRN) and Christian-Albrechts-University of Kiel (CAU).

#### Data availability statement

The model data that support the findings of this study will be openly available on Ocean-Rep - GEOMAR Repository (oceanrep.geomar.de) shortly after the publication. The source code of TMM is available at https://doi.org/10.5281/zenodo.1246300. The code of our calibration framework are available from the corresponding author upon reasonable request.

## 4. Simulated future trends in marine nitrogen fixation are sensitive to model iron implementation

This chapter is based on the manuscript "Simulated future trends in marine nitrogen fixation are sensitive to model iron implementation" submitted to the journal Global Biogeochemical Cycles in October 2020.

#### 4.1. abstract

Biological nitrogen fixation is an important oceanic nitrogen source, potentially stabilizing marine fertility in an increasingly stratified and nutrient-depleted ocean. We estimate future changes to the marine nitrogen cycle from three biogeochemical models, calibrated against observed nutrients and oxygen data, using a business-as-usual RCP8.5 atmospheric CO<sub>2</sub> trajectory to the year 2100. The first model neglects explicit iron effects on biology (NoFe), the second utilizes prescribed, seasonally-cyclic iron concentrations and associated limitation factors (FeMask), and the third contains a fully dynamic iron cycle (FeDyn). Global marine nitrogen fixation increases by 71.1% with respect to the preindustrial value by the year 2100 in NoFe, while it remains stable (0.7% decrease in Fe-Mask and 0.3% increase in FeDyn) in iron models. In NoFe, warming-enhanced primary production outpaces increased zooplankton grazing to increase export production, which decreases water column oxygen concentrations and expands oxygen deficient zones . Enhanced denitrification in these oxygen deficient zones decreases the downstream nitrate supply with respect to phosphate, promoting the fixation of nitrogen across the tropical ocean. In models with iron, the additional bottom-up control of phytoplankton growth provided by iron limitation in the upwelling regions decreases the regional export production, decreasing the size of the oxygen deficient zone, and mitigates the downstream nitrogen fixation response under global warming.

#### 4.2. Introduction

Marine ecosystems provide crucial services to humanity, yet are under pressure due to climate change. Most global earth system models agree on a future ocean with increasing stratification and decreasing nutrient re-supply to the surface mixed layer, particularly in low and mid-latitudes (Cabré et al., 2015). How marine ecosystems will respond to these changing conditions is uncertain. The growth of phytoplankton is controlled by

nutrient and light limitation, temperature, and grazing pressure Parsons et al. (1984). Differences in model parameterization can lead to wide-ranging ecosystem responses to transient forcing (e.g., Taucher and Oschlies, 2011; Kwiatkowski et al., 2014). Climate model intercomparison reveals that while ocean surface temperature, pH and global oxygen (O<sub>2</sub>) content show relatively similar trends across models, the change in simulated global Net Primary Production (NPP) by the end of the 21st century varies widely both in magnitude and sign, with particularly large differences in low latitude NPP trend (Bopp et al., 2013). This spread in model projections of NPP is attributed to differences in parameterization of surface nutrient re-supply pathways, e.g., the remineralization of organic particles, zooplankton grazing and microbial loop recycling, in addition to the strength of nutrient limitation realized in the different models (Laufkötter et al., 2015; Landolfi et al., 2017).

Nitrogen (N) is an important nutrient in the marine ecosystem. In an increasingly nutrientlimited surface ocean, processes that could add "new" N are potentially important for the maintenance of ocean fertility. The biological fixation of dinitrogen (N<sub>2</sub>) by diazotrophs is a major natural source of fixed N to the ocean (140  $TgNyr^{-1}$ , Gruber and Galloway, 2008; Karl et al., 2002). Newly fixed N becomes biologically available to non-diazotrophic phytoplankton (Mulholland, 2007). Earth system models produce a range of N<sub>2</sub> fixation trends under climate change, which arise from differences in model structure (Wrightson and Tagliabue, 2020).

Improving the constraints on future  $N_2$  fixation trends requires an improved understanding of the processes critical for modern ocean  $N_2$  fixation. Spatial patterns of  $N_2$  fixation are thought to be determined by three key factors: denitrification and its impact on macronutrient ratios, iron availability, and grazing pressure. The denitrification (and anammox) that occurs in low oxygen environments (such as in tropical upwelling zones) changes the relative availability of nitrate (NO<sub>3</sub>) and phosphate (PO<sub>4</sub>) in the surface ocean downstream (Deutsch et al., 2007; Weber and Deutsch, 2014). The resulting stoichiometric excess phosphate (positive P<sup>\*</sup>, defined as PO<sub>4</sub> - NO<sub>3</sub>/16, Deutsch et al., 2007) might provide diazotrophs with a competitive advantage relative to other types of phytoplankton. This competitive advantage is hypothesized to explain why rates of N<sub>2</sub> fixation are enhanced in the NO<sub>3</sub>-depleted low latitude surface ocean (Tyrrell, 1999; Deutsch et al., 2007; Garcia et al., 2015).

At the same time, iron (Fe) is essential for all phytoplankton as it is needed to form a range of cellular enzymes that are associated with photosynthesis and respiration. However, diazotrophs additionally require iron for producing the  $N_2$  fixation enzyme nitrogenase (Raven, 1988; Kustka et al., 2003b; Tagliabue et al., 2017; Zehr and Capone, 2020), causing diazotroph growth to be particularly sensitive to a lack of iron. For diazotrophs, the supply of iron is so essential that its availability helps to define the biogeography of  $N_2$  fixation (Ward et al., 2013; Moore et al., 2009; Schlosser et al., 2014). Iron availability and denitrification might also be linked in the ocean. Biological productivity is limited by iron availability over one-third of the surface ocean, including the Eastern Tropical Pacific (ETP) upwelling zones (Moore et al., 2013). This limitation can impact particulate organic matter export (Martin, 1990). Therefore, a change in iron availability in a highly productive upwelling zone might alter particulate organic matter export and thereby influence the extent and intensity of oxygen deficient zones (ODZs) and of denitrification. In their modeling study, Buchanan et al. (2019) showed that an increase in iron availability in the ETP can lead to a larger ODZ, and a downstream enhancement of diazotroph biomass and  $N_2$  fixation. They inferred that this process could have contributed an additional 7-16 *ppm* of carbon dioxide (CO<sub>2</sub>) sequestration in the tropical ocean in the Last Glacial Maximum (Buchanan et al., 2019).

In addition to iron availability and denitrification, top-down control of diazotrophy by zooplankton grazing might also influence the global patterns of N<sub>2</sub> fixation (Wang et al., 2019). The growth rate of diazotrophs is generally lower than that of other phytoplankton (Edwards et al., 2015; Fu et al., 2014), as N<sub>2</sub> fixation is 25% more energetically costly compared to direct uptake of NO<sub>3</sub> from the surrounding water (Knapp et al., 2012). Yao et al. (2019) recently demonstrated in a model study that for high zooplankton grazing rates, diazotroph biomass can be subject to top-down control. Strong zooplankton grazing was found to prevent diazotrophs inhabiting the ETP, despite high P<sup>\*</sup> and when iron was not limiting in their model (Yao et al., 2019).

Here we examine what controls the global  $N_2$  fixation response to global warming in three marine biogeochemical models with different representations of the marine iron cycle, all of which perform similarly with respect to reproducing observed NO<sub>3</sub>, PO<sub>4</sub> and O<sub>2</sub> distribution (Yao et al., 2019). All three models were individually subjected to parameter optimization and their preindustrial steady-states were assessed in Yao et al. (2019). We first describe the models and the new transient global-warming experiments. Then, we compare the different simulated responses of N<sub>2</sub> fixation to global warming. Next, we examine the impact of ODZ development and denitrification on N<sub>2</sub> fixation. We link the development of ODZs to in situ export production trends via the different assumptions about nutrient pathways in the different models, and make some recommendations for further model improvements.

#### 4.3. Method

#### 4.3.1. model description

In this study, we use the University of Victoria Earth System Climate Model (UVic ESCM, version 2.9; Weaver et al., 2001; Eby et al., 2013). The UVic ESCM is an earth system model of intermediate complexity. It contains four components: a simple one-layer

atmospheric model, a terrestrial model (Meissner et al., 2003), a sea-ice model and a three-dimensional ocean general circulation model. All components have a horizontal grid resolution of 3.6 degree in meridional and 1.8 degree in zonal direction. The atmospheric model is based on an energy-moisture balance scheme (Fanning and Weaver, 1996), which calculates heat and water fluxes between the atmosphere and the ocean, land and sea ice dynamically, while using prescribed (NCAR/NCEP) winds. A thermodynamic sea-ice model (Bitz and Lipscomb, 1999) is applied, while constant continental ice sheets are prescribed. The ocean component is the Modular Ocean Model 2 (MOM2). It has 19 vertical layers with increasing thickness from 50 m near the surface to 500 m in the deep ocean. The advection scheme is a non-linear second-order Flux-Corrected Transport scheme (Weaver and Eby, 1997).

#### 4.3.2. biogeochemical models and parameterization

We apply three variants of the marine biogeochemical component in the UVicESCM, which differ in their treatment of the iron cycle. All of them are nutrient-phytoplankton-zooplankton-detritus (NPZD) models with two types of phytoplankton, ordinary phytoplankton and diazotrophs, one zooplankton, and detritus. In the models, the growth of phytoplankton is governed by nutrients, light and temperature. In all three models, diazotrophs grow slower than ordinary phytoplankton but can also fix  $N_2$  to meet their N demand when  $NO_3$  is limiting.

The first model contains no representation of the iron cycle (NoFe, created by disabling iron limitation in model presented by Keller et al., 2012). The second model variant makes use of a prescribed seasonally periodic iron concentration mask (FeMask; Keller et al., 2012), which is re-gridded from the output of the Biology Light Iron Nutrient and Gases model (BLING; Galbraith et al., 2010). The third variant includes a fully dynamic iron cycle (FeDyn; Nickelsen et al., 2015), which resolves major sources and sinks of iron in the ocean, e.g., dust deposition, benthic release, particle scavenging, colloid formation, remineralization as well as hydrothermal release (Yao et al., 2019). In the models that include iron (FeMask and FeDyn), both diazotroph and ordinary phytoplankton growth rates are reduced when iron is limiting, in addition to the usual controlling factors, e.g., macro-nutrient(s), light and temperature. In the model without iron (NoFe), the growth rates of diazotrophs and ordinary phytoplankton do not depend on the iron concentration. Details, including equations, of each model configuration can be found in Nickelsen et al. (2015, FeDyn) and Keller et al. (2012, FeMask and NoFe).

Biogeochemical parameters of all three model variants were calibrated individually for each model using present-day global climatologies of NO<sub>3</sub>, PO<sub>4</sub> and O<sub>2</sub> (World Ocean Atlas 2013; Garcia et al., 2013b,a). The details of the calibration can be found in Yao et al. (2019), which produces three differently structured models with similar global misfits to NO<sub>3</sub>, PO<sub>4</sub> and O<sub>2</sub>. We apply the optimal parameter sets from that steady-state



**Figure 4.1.:** Simulated change in 5-yearly averaged global marine NPP (a) and diazotroph primary production (b) between 1800 and 2100 for all models: NoFe (red), FeMask (green) and FeDyn (blue).

calibration to the transient global warming runs in this study (table B.1).

#### 4.3.3. model experiments

All models are integrated for ten thousand years with pre-industrial (year 1800) climate boundary conditions to reach equilibrium. A prescribed historical and business-as-usual RCP8.5 atmospheric CO<sub>2</sub> concentration forcing (van Vuuren et al., 2011) is then applied from year 1800 to 2100, but with climatological seasonally cycling NCAR/NCEP wind fields. The preindustrial physical circulation and circulation responses are identical in all three model configurations. Atmospheric iron deposition flux (Luo et al., 2008), which is a preindustrial climatological estimate of monthly iron deposition, is not changed during the transient experiment in FeDyn. We do not include atmospheric N deposition in this study, which was found to be of minor importance in an earlier study by Landolfi et al. (2017).

#### 4.4. Results and Discussion

### 4.4.1. Trends in global NPP and diazotroph primary production show inter-model differences

Global total primary production in year 1800 is 74.8 (NoFe), 49.4 (FeMask), and 50.4 (FeDyn)  $GtCyr^{-1}$ . In all simulations, NPP decreases with time (figure 4.1a), in agreement with the trends of most CMIP5 models (Laufkötter et al., 2015; Bopp et al., 2013). In the year 2100, the largest absolute decline is found in NoFe (5.6  $GtCyr^{-1}$ ) and the

largest percentage decline in FeMask (9.16%).

Global diazotroph primary production in year 1800 is 0.6 (NoFe), 0.3 (FeMask), and 0.4 (FeDyn)  $Gt Cyr^{-1}$  (figure 4.1b). In contrast to global NPP, diazotroph primary production increases in all simulations (by 0.02 to 1.0  $Gt Cyr^{-1}$  between 1800 and 2100; figure 4.1b). The largest percentage increase is found in NoFe (177%) and the smallest in FeDyn (4.2%). When compared with total NPP, the relative importance of diazotroph primary production to total NPP increases between 1800 to 2100. In 1800, diazotroph primary production contributes 0.76% (NoFe), 0.57% (FeMask), and 0.75% (FeDyn) to global NPP and in 2100 it increases to 2.3% of global NPP in NoFe, 1.1% in FeMask and 0.82% in FeDyn.

4.4.1.1. Trends in N<sub>2</sub>-fixation diverge between models

The simulated pre-industrial N<sub>2</sub> fixation rates are between 30.3 and 54.1  $TgNyr^{-1}$  (figure 4.2.a). The calibrated NoFe has the smallest preindustrial N<sub>2</sub> fixation rate, which is also much smaller than earlier, hand-tuned versions of the UVicESCM that applied different parameter values and also did not include iron (UVicESCM version 2.7, 316  $TgNyr^{-1}$ , Riche and Christian, 2018; UVicESCM version 2.8 & 2.9, 150  $TgNyr^{-1}$ , Oschlies et al., 2019 and Keller et al., 2012; UVic ESCM version 2.9 with additonal structual changes including benthic denitrification,  $110 - 126 TgNyr^{-1}$ , Landolfi et al., 2017). Applying a dynamic iron cycle (FeDyn) in the model results in the largest preindustrial  $N_2$  fixation rate among the three models. However, even the FeDyn  $N_2$  fixation rate is in the lower range of observations and model estimates (between 89 and 213  $TgNyr^{-1}$ ; Landolfi et al., 2018). Note, that independent estimates of the  $N_2$  fixation rate were not used as an explicit constraint in the calibration of Yao et al. (2019). Furthermore, none of the models include benthic denitrification, which could increase global total denitrification and enhance global N<sub>2</sub> fixation (Weber and Deutsch, 2014). The simulated evolution of N<sub>2</sub> fixation in response to global warming differs considerably between simulations (figure 4.2.a&b). While N<sub>2</sub> fixation increases by 21.6  $TgNyr^{-1}$  (71%) in NoFe between 1800 and 2100, fixation rates remain almost unaffected by global warming in FeMask and FeDyn. Landolfi et al. (2017) reported an increase of 0.3% in N<sub>2</sub> fixation due to the increasing PO<sub>4</sub> limitation for the same simulated period in a version of UVicESCM with prescribed iron concentrations. Another similar result is found in a model comparison study by Wrightson and Tagliabue (2020), where the only model that had a strong increase in N<sub>2</sub> fixation (58  $TgNyr^{-1}$ ) between years 1800 and 2100 is a model with neither iron nor phosphate limitation on (diagnosed) N2 fixation. However, two models with neither iron nor explicit diazotrophs in their study showed large decreases in N<sub>2</sub> fixation (38  $-50.1 TgNyr^{-1}$ ) over this time period. Most of their models with representation of iron limitation project decreasing N<sub>2</sub> fixation (up to 23.2  $TgNyr^{-1}$ ; Wrightson and Tagliabue, 2020) while we find the  $N_2$  fixation trend in FeMask and FeDyn to be almost unaffected.


**Figure 4.2.:** Simulated running 5-year-averaged global  $N_2$ -fixation rates (a), their temporal changes (b), temporal changes of denitrification (c) and the changes in ODZ volume (d, change in the percentage of the total ocean volume) between 1800 and 2100. Note: we define ODZ as regions where  $O_2$  concentration is below 5 mmol  $m^{-3}$ .



**Figure 4.3.:** Surface (0-130 *m*) Surface N cycle (number before + or - sign) and its trend (number after + or - sign) in the Eastern Tropical Pacific (Longitude: 115W:60W, Latitude: 8S:20N) between years 1800 and 2100. The unit for inventory is TgC and for fluxes is  $TgCyr^{-1}$ . The thickness of the arrow is approximately proportional to the respective flux. Colors denote different models: NoFe (red), FeMask (green), and FeDyn (blue).

4.4.2. The trends of the Eastern Tropical Pacific ecosystem response to warming differ

We next look at differences in the model responses to global warming in the ETP, due to its apparently large impact on global  $N_2$  fixation trends (discussed below).

As the temperature increases, NPP, remineralization, microbial loop recycling, and other temperature-sensitive fluxes accelerate accordingly (figure 4.3). In all models, the surface ocean nutrient recycling pathways through enhanced and shallower detrital remineralization and the enhanced microbial loop gain importance under warming. Also, the increase in phytoplankton potential growth rate is greater than the increase in zooplankton potential grazing rate (which is capped above 20 °C (Keller et al., 2012) in all models. Hence, total plankton biomass (ordinary phytoplankton, diazotrophs and zooplankton) increases in all models in the ETP.

In the NoFe model, light limitation due to phytoplankton self-shading is the dominant

control on phytoplankton growth in the ETP (figure B.1). Grazing control on the biomass is also stronger compared to the other models (NoFe has the highest optimized maximum grazing rate at 0°C, see table B.1). With ocean warming, the increase in phytoplankton potential growth rate is larger than the increase in zooplankton potential grazing rate in the ETP. This creates a 23% higher total plankton biomass in the region, which leads to a 35% higher net detritus production (sloppy feeding of zooplankton plus plankton mortality and lysis, minus zooplankton grazing on detritus, figure 4.3). Warming also enhances detrital remineralization, however, a 41% increase in detrital remineralization (figure 4.3) is not enough to keep up with the 401.0  $TgCyr^{-1}$  increase in net detritus production. Therefore, export production in the ETP increases 15% (figure 4.3), contributing to an ODZ volume increase.

Phytoplankton are nutrient limited in the ETP (by Fe and NO<sub>3</sub>, figure B.1) in both models with iron, which makes their production more sensitive to decreasing nutrient resupply when water stratification intensifies. In FeDyn and FeMask, euphotic zone detrital remineralization accounts for around 23% of the NPP compared to 17.2% in NoFe (figure 4.3). Warming in the ETP increases production in FeDyn and FeMask, but also particle remineralization and microbial loop nutrient recycling. The net effect is a decrease in export production in both models (5.1% for FeDyn and 0.8% for FeMask, figure 4.3). The optimized microbial loop constant is larger in FeDyn compared to FeMask (table B.1), and this causes a larger decline in export production in the ETP in this model (figure 4.3). In both iron models there is a reduction in ODZ volume in the ETP (figure B.2.h&i). This has global consequences for N-cycle coupling.

# 4.4.3. The decoupling of $N_2$ -fixation and diazotroph primary production in models with iron

In the NoFe model, we find increasing  $N_2$  fixation correlates with an increase in diazotroph primary production, which is intuitive as diazotrophs fix  $N_2$  to support their growth when NO<sub>3</sub> is in short supply. In the tropics (30S:30N), global warming increases the temperature-dependent phytoplankton (both diazotrophs and ordinary phytoplankton) potential growth rate and the detritus remineralization rate by around 40% from preindustrial to year 2100, which accelerates the surface nutrient cycle and promotes phytoplankton growth in oligotrophic waters. At the same time, elevated temperatures in the tropical upwelling zones increase export production (figure 4.4.a) and particle remineralization, which decreases water column oxygen and increases denitrification, producing increasingly elevated P<sup>\*</sup> (figure 4.5) conditions downstream that favor diazotrophs and promote N<sub>2</sub> fixation.

However, in FeMask and FeDyn, diazotroph primary production increases over time but  $N_2$  fixation does not. This is because diazotroph primary production and  $N_2$ -fixation can be decoupled (Knapp et al., 2012). In our models, diazotrophs can utilize NO<sub>3</sub> when it



**Figure 4.4.:** Change in annual export production at 130 *m* depth (years 2100 minus 1800; from left to right, NoFe, FeMask, FeDyn). Unit for the color scale is  $gCm^{-2}yr^{-1}$ . The solid thick contour line denotes 0, and the solid (dashed) contour denotes positive (negative), the interval is 8  $gCm^{-2}yr^{-1}$ . The dashed red rectangles represent the Eastern Tropical Pacific, Atlantic and Bay of Bengal regions. Note: the Caribbean region is excluded from the calculations for the box in the Eastern Tropical Pacific.



**Figure 4.5.:** The change of annually averaged  $P^*$  concentration  $(mmol m^{-3})$  at 350 m depth between years 1800 and 2100. The solid thick contour line denotes 0, and the solid (dashed) contour denotes positive (negative) values, the interval of contour is 0.1 mmol m<sup>-3</sup>. While PO<sub>4</sub> and NO<sub>3</sub> were used as data constraints in the model calibration study (Yao et al., 2019), P<sup>\*</sup> itself was not.

is available. Similar to NoFe, surface nutrient cycling is enhanced by warming in the oligotrophic regions for FeMask and FeDyn (including diazotroph primary production). However, persistent iron limitation in the tropical upwelling zones (figure B.1) mitigates the particle export response to regional warming (figure 4.4; the net export in the three ODZ regions declines by  $66.5 TgCyr^{-1}$  for FeMask and  $69.0 TgCyr^{-1}$  for FeDyn by year 2100). In line with these trends in particle export, a net contraction of the ODZ volume occurs in both models, which reduces total denitrification (figure 4.2.c&d). This decline in denitrification over time results in reduced P\* (figure 4.5) for the downstream surface ocean, which tends to suppress the temperature-induced growth rate increase of diazotrophs and the associated N<sub>2</sub> fixation in the stratified, oligotrophic regions. The strength of this effect is subject to model parameter differences. How sensitive the modelled denitrification response is to fluctuations in water column oxygen depends on the prescribed O:N ratio, which was determined in the parameter calibrations (Yao et al., 2019) and varies between 9.54 and 10.30 among the models (table B.1).

#### 4.4.4. FeDyn v.s. FeMask

We find no significant differences in the global  $N_2$  fixation trends between FeDyn and FeMask models over the time period simulated. This may be due to the models having relative similar parameter values, or to compensating effects for iron concentration in the dynamic iron model (e.g. iron scavenging and remineralization). The iron scavenging and colloid formation can quickly remove iron from the dissolved phase, which keeps the change in iron concentration small in the ETP (0.4% increase). This small change is not able to qualitatively change the model response in export production in the upwelling zone, hence there is little impact on  $N_2$  fixation trend. However, we do see regional differences in other quantities, such as export production trends in the Southern Ocean, which suggest that differences may be significant in other respects. In addition, FeMask has prescribed surface iron concentrations, but surface iron concentrations are expected to increase with climate warming (Fu et al., 2016). These dynamic changes are not represented in FeMask.

However, FeDyn also has deficiencies worth mentioning, with the most prominent being its poor skill in reproducing observed iron concentrations (Yao et al., 2019). Considering the global averaged 0.57-year turnover time in FeDyn (compare to other models ranging from a few years to more than 500 years; Tagliabue et al., 2016), the dissolved iron concentration depends strongly on local iron sources, e.g., dust deposition, sediment releases, and hydro-thermal activity. Sediment release is the largest iron source in our model and better parameterization might be beneficial. A recent laboratory study showed that certain types of diatoms can reduce their iron requirement under higher temperatures and maintain their growth rate (Jabre and Bertrand, 2020). This finding suggests a more flexible, even a temperature dependent iron demand parameterization, might be appropriate to include in the future. Likewise, the Fe dust deposition is projected to increase in the future



**Figure 4.6.:** A comparison of two different model responses to tropical warming and water column stratification, with (right side) and without (left side) explicit consideration of iron limitation. The upwelling ecosystem is controlled by a combination of top-down grazing and light limitation when iron does not limit primary production (left side). As the temperature increases, upwelling is reduced ("-" in the center of the figure). In the scenario not considering iron (left part of the figure) primary production increases faster than grazing pressure, which leads to increased particle export (+) and higher remineralization (+). Higher temperatures and increased remineralization intensify water column ODZs (+) and associated denitrification (+). Elevated denitrification leads to higher N<sub>2</sub> fixation, and an increase in diazotroph primary production response to warming in the upwelling ecosystem. Export production declines (-), and the ODZ volume also shrinks (-). More oxygen in the water column leads to less denitrification (-) and reduced P<sup>\*</sup>, which suppresses N<sub>2</sub> fixation downstream despite temperature-driven increases in diazotroph primary production.

(Hamilton et al., 2020), which we did not include in FeDyn for this study.

Wrightson and Tagliabue (2020) also compiled a list of non-iron model related potentially significant drivers for  $N_2$  fixation, e.g.,  $CO_2$  fertilization on diazotrophs, and an upper thermal threshold for optimal growth temperature for phytoplankton, which are not yet implemented in our model and should be explored in the future.

## 4.5. Conclusions

The results presented here demonstrate divergent trends in future global  $N_2$  fixation and denitrification that are determined by the representation of bottom-up and top-down phy-

toplankton growth limitation in tropical upwelling regions (schematic in figure 4.6). When iron limitation is neglected, warming induces positive feedbacks of enhanced regional export production and denitrification that triggers enhanced  $N_2$  fixation across the central and western basin. Iron limitation mitigates the export production response to warming and even produces a net decline in denitrification in the ETP, leading to essentially no net change in global marine  $N_2$  fixation. That the  $N_2$  fixation responses diverge between the three models tested, despite all models simulating an increase (albeit, of variable magnitude) in diazotroph primary production illustrates the potential for a global decoupling of these two seemingly co-dependent factors.

That the N<sub>2</sub> fixation trends hinge upon the treatment of iron limitation in the tropical upwelling zones highlights the importance of model structure. This finding is particularly salient considering all models have similar global misfits with respect to state variables NO<sub>3</sub>, PO<sub>4</sub>, and O<sub>2</sub> in steady-state, and suggests flux and rate data may be just as (or even more) relevant for constraining models with calibration techniques. Our finding of a strong connection between global N<sub>2</sub> fixation rates and iron limitation in upwelling zones furthermore suggests a more inter-connected view between two controlling factors in N<sub>2</sub> fixation, iron limitation and denitrification. From a global view, spatial patterns of N<sub>2</sub> fixation are not only controlled by iron availability in oligotrophic regions (as proposed by Ward et al., 2013), but also by the iron limitation upstream, which regulates the denitrification (and hence, P<sup>\*</sup>) pattern. Lastly, our study demonstrates a role of temperature in determining N<sub>2</sub> fixation patterns.

In our earlier manuscript we demonstrated a correlation between denitrification and  $N_2$  fixation as well as globally balanced rates in all model steady states (Yao et al., 2019). However, this linkage is reduced when climate warming is applied. As temperature increases, temperature-sensitive rates regulating bottom-up and top-down control on export production respond differently. The net result is a decreasing sensitivity of  $N_2$  fixation to denitrification (seen particularly in the models with iron) that can to lead to at least a temporary imbalance in global rates. Furthermore, there is growing evidence that non-primary producers may play a roll in  $N_2$  fixation (Moisander et al., 2017), and that  $N_2$  fixation in key cyanobacterial diazotophs may not be stimulated by the lack of NO<sub>3</sub> (Mills et al., 2020). Both of these factors can contribute further to the decoupling between  $N_2$  fixation, denitrification and diazotrophs primary production. The long-term implications of potential  $N_2$  fixation and denitrification imbalance would be an interesting topic worthy of further study.

#### Acknowledgments

The authors would like to acknowledge computer resources from Christian-Albrechts-University of Kiel (CAU) and GEOMAR Helmholtz Centre for Ocean Research, Kiel. This work is supported by the Helmholtz Research School for Ocean System Science and Technology (HOSST, Germany) at GEOMAR Helmholtz Centre for Ocean Research Kiel (VH-KO-601). All authors are grateful for stimulating discussion within the Biogeochemical Modeling and Marine Chemistry research units at GEOMAR, Dalhousie University, and University of Liverpool.

Data Availability Statement

Model code and model data used in writing this manuscript are available on OceanRep - GEOMAR Repository (https://oceanrep.geomar.de/ link in preparation).

# 5. Conclusions

Iron plays an important role in global marine biogeochemistry, yet the benefits of implementing a complex iron model into a biogeochemical model have been unclear. Disregarding the extra computing expenses that comes with increasing complexity, the tuning of parameters in relatively less complex biogeochemical models is already a challenge for modellers with a limited pool of independent data sets as constraints. How wise is it to introduce more complexity by adding an iron cycle into a biogeochemical model?

In this thesis, I aimed to elucidate the impact of iron models with different complexities for global biogeochemical cycles, both in steady state and with transient climate change. I selected two representations of iron cycling in the UVic 2.9; one prescribes iron concentrations (FeMask), and another contains a complex iron cycle (FeDyn). A third version of the model ignores the impact of iron on phytoplankton growth (NoFe). I calibrated each variation of the model against the same set of global observations of nitrate, phosphate, and oxygen concentrations and found differences in the resulting surface nutrient pathways in the pre-industrial climate among models. I also performed transient simulations with all three calibrated models and found diverse nitrogen fixation responses towards warming among models.

In the preparation of model calibration, I performed a sensitivity test of the UVic 2.9 in Chapter 2. A literature search for parameters in the UVic 2.9 shows that some of the values have not been updated since the earlier release of KMBM (e.g., the maximum growth rate constant at 0°C, Schmittner et al., 2005) and some refer to measurements which date back more than 30 years (e.g., the iron scavenging parameters are from Honeyman et al., 1988). Some parameters cannot be directly measured in the field (e.g., the microbial loop recycling rate, one parameter that describes the whole microbial community and its impact on the organic matter in surface ocean ecosystem). Despite this, and the fact that most parameters are hand-tuned against surface nutrient fields, the correlation from the model version by Nickelsen et al. (2015) is around 0.85 to 0.9 against whole ocean nitrate, phosphate and oxygen distributions. However, a more comprehensive sensitivity test of the parameters shows potential for a few parameters to achieve better model fitness (up to 0.95). This indicates that the model ability to reproduce observational data will benefit from parameter calibration.

When comparing the full dynamic iron cycle model against the current iron observational data, it is found that parameters are insensitive for the model misfit. This shows how much potential there is for improvement in the field of iron modelling. This might also indicate

needs for improvements in data coverage and spacial interpolation. Models with intermediate complexity are not designed to simulate local or regional spatial scales, therefore greater spatial (and temporal) iron observational data are needed. Meanwhile, isotope data also help to improve understanding of fluxes and pathways in the iron cycle.

In the follow-up model calibration study, an 8.6% improvement of global integrated misfit against nutrient and oxygen climatologies compared to the previously hand-tuned version (Nickelsen et al., 2015) was achieved. Furthermore, in this study, through the comparison between the three calibrated versions of UVic 2.9 it is shown that the model with a full dynamic iron cycle has a marginal (up to 4.8%) better skill of reproducing biogeo-chemical distributions in the pre-industrial, steady-state climate compared to the others. The model with a full dynamic iron cycle also has advantages over the other two models with respect to independent metrics, e.g., net primary production and oxygen deficit zone volume. As the calibration tries to fit the models to the observational data as close as it can, maybe it is not that surprising that biogeochemical models coupled with the same circulation could end up with similar model misfits. However, due to the differences in model complexity, all models have different surface nutrient pathways, which represents very different pictures of how marine ecosystems function. These results offer the question whether those different nutrient pathways can produce different responses to climate change, despite similar performance with respect to the applied metrics at steady-state.

From the calibrations performed for this thesis, I conclude that hand-tuning may be not sufficient for handling more complex biogeochemical models. The calibration framework applied in this thesis has shown promising model misfit improvements, yet was limited to five or six parameters per model (only a small portion of the total number of parameters). Also, biogeochemical parameters are only part of what determines biogeochemical distributions. Calibration of the biogeochemical parameters may not be enough to overcome all bias in the ocean circulation model (e.g., a too weak formation and northward propagation of Antarctic Intermediate Waters and the shift of the oxygen deficit zone from the Arabian Sea to the Bay of Bengal). This means that the resulting biogeochemical parameter values may reflect circulation dependency. However, in the study by Kriest et al. (2020), they show that some biogeochemical parameters can be transferable to another model with similar circulation.

In Chapter 4, I compared transient simulations (over the 21st century) from the three calibrated versions of UVic 2.9. Simulated nutrient pathway differences lead to similar global trends of NPP and export production from year 1800 to 2100 and there are no significant differences in marine carbon uptake between models (ranging from 2017 to 2040 *GtC* accumulated for the whole simulated period). However, they produce different regional export production responses to warming in the upwelling regions, which creates disagreement in the change of the oxygen deficit zone volume. As the oxygen deficit zone is where denitrification occurs in the model, this disagreement leads to different nitrogento-phosphorus supply ratios downstream. Hence, different trends in nitrogen fixation in

the downstream oligotrophic regions are projected. This shows that iron not only impacts the marine nitrogen cycle by its availability in oligotrophic regions, where nitrogen fixation occurs, but also by regulating export production in the upwelling regions. This advances our understanding of how iron influences the global marine nitrogen cycle.

All model versions used in this thesis have similar global misfits against observed nitrate, phosphate and oxygen fields in steady-state, yet present different trends in the transient. I conclude that iron models are an important component of global biogeochemistry, despite their current poor performance at reproducing observed iron concentrations. Given the different trends are related to nitrogen fixation, I conclude that the simulation of iron is important for studies on future ocean fertility (e.g., fishery management). However, I find no significant changes in the ocean carbon uptake related to the different implementations of iron in model. This indicates that iron may not be the bottleneck of the soft tissue carbon pump under the high emission scenario until year 2100 studied here.

When comparing models that include iron, transient simulations are qualitatively similar in global model response of nitrogen fixation, which might suggest that a model with a prescribed iron concentration field is sufficient for projecting nitrogen cycle trends over the 21st century. Less complexity means lower computational demands and less parameter uncertainty, which may make a more simplistic iron implementation more desirable. Nevertheless, the model with a prescribed iron concentration field has different trends in net primary production and export production in the upwelling regions comparing to the model with a full dynamic iron cycle. This may come from unrealistic iron availability in the model with prescribed iron concentrations. This approximation might lead to more strongly diverging model differences over longer period simulations.

The models presented in this thesis show different surface nutrient pathways after calibration against global distributions of nitrogen, phosphorus, and oxygen. This indicates that the existing concentration data used for model calibration may not be enough to unambiguously constrain the fluxes within the marine ecosystem, which creates uncertainty in the system behavior in the face of change. Biogeochemical flux data are rare but still beneficial to be incorporated as criteria in future model calibration studies.

### 5.1. Outlook

In the thesis, I calibrated only a portion of the total parameters in the UVic 2.9. I suggest a second calibration with another selection of parameters in order to see if a set of newly calibrated models behave the same way as the models in this thesis under the same forcing settings. The current calibration of the iron cycle is focused on the iron scavenging and the phytoplankton iron uptake rates, which are relatively insensitive for the model misfit against current iron observational data sets. In the next calibration, I suggest testing a large range of iron source fluxes, which are sensitive to model misfit against iron observations (Pasquier and Holzer, 2017), so that the iron observational data can be assimilated. This may provide insight on how to improve our iron model.

I found in this thesis that a more complex iron model has a marginally better performance at reproducing the observational patterns of nitrate, phosphate, and oxygen. However, the most complex model in this thesis is still a simplification of the iron cycle. Would an additional increase in complexity (e.g., inclusion of a more sophisticated ligand representation) bring additional gain in performance? Which is the bottleneck in model performance, model complexity or model calibration capacity? Considering that the misfit differences between calibrated models are smaller than the differences for the same model before and after the calibration, I speculate that the latter is more likely.

In the model examined in this thesis, phytoplankton have a fixed ratio between iron and nitrate for its biomass, which is an assumption inherited from Redfield stochiometry. However, in reality, phytoplankton can take up more iron than needed when the environment is iron replete (Iwade et al., 2006; Chen et al., 2011; Shire and Kustka, 2015; Twining et al., 2021) and phytoplankton can recycle biogenic iron from their own cells to support growth under iron deficiency (Louropoulou et al., 2019). Not like nitrogen nor phosphorus, which are the building blocks of organic material, iron appears to be more like the building crew, or a catalyst of production. It seems that the iron quota in phytoplankton may be strongly influenced by the environment. If that is the case, then the impact of projected increasing iron availability in the future (Hamilton et al., 2020) may be amended by a future enhanced iron quota in phytoplankton. I suggest including a flexible iron quota in the next version of the iron model.

It seems that iron scavenging cannot be constrained by the current observational dissolved iron concentration data. The dissolved iron concentration is a result of interplay among ligands, iron sources, and sinks. The sinks include scavenging and biological uptake. The sources such as sedimentary flux are hard to estimate without a prior estimation of the scavenging strength and the biological uptake flux. Hence I suggest measuring or estimating iron scavenging in more isolated environments, such as laboratory experiments or in the deep ocean hydrothermal plumes, where biological processes are absent.

From the 1800s to year 2100, models with simple and complex iron implementations have similar trends in nitrogen fixation. However, due to the differences in nutrient pathways in the Eastern Tropical Pacific region, the differences in detritus export have increased (See Fig. 3.2) and may lead to larger differences of nutrient inventory in the surface ocean, and hence impact global marine productivity (See Fig. 4.1, the models have different rates of decrease in global net primary production). Would this lead to different sensitivities to further warming? As the net community production will bounce back under the scenario of zero carbon dioxide emission after year 2300 (Oschlies et al., 2019), would the implementation of iron in that model impact the recovery time? Those questions can be

investigated through a set of simulations under identical forcing with a longer integration time.

Lastly, this thesis emphasizes that the tropical upwelling regions, such as the Eastern Tropical Pacific, can be one of the key factors to determine the ocean nitrogen inventory with climate change. Hence, I suggest improving our understanding of the marine ecosystems in those regions and their responses towards warming (e.g., export production change and oxygen deficient zone volume evolution).

# Bibliography

- Anderson, T. R. (2005). Plankton functional type modelling: running before we can walk? *Journal of Plankton Research*, 27(11):1073–1081.
- Archer, D. E. and Johnson, K. (2000). A model of the iron cycle in the ocean. *Global Biogeochemical Cycles*, 14(1):269–279.
- Aumont, O. and Bopp, L. (2006). Globalizing results from ocean in situ iron fertilization studies. *Global Biogeochemical Cycles*, 20(2):n/a–n/a.
- Beghoura, H., Gorgues, T., Aumont, O., Planquette, H., Tagliabue, A., and Auger, P. (2019). Impact of Inorganic Particles of Sedimentary Origin on Global Dissolved Iron and Phytoplankton Distribution. *Journal of Geophysical Research: Oceans*, 124(12):8626–8646.
- Berelson, W. M. (2001). Particle settling rates increase with depth in the ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(1-3):237–251.
- Bitz, C. M. and Lipscomb, W. H. (1999). An energy-conserving thermodynamic model of sea ice. *Journal of Geophysical Research: Oceans*, 104(C7):15669–15677.
- Bloesch, J. and Burns, N. M. (1980). A critical review of sedimentation trap technique. *Schweizerische Zeitschrift für Hydrologie*, 42(1):15–55.
- Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J., and Vichi, M. (2013). Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. *Biogeosciences*, 10(10):6225–6245.
- Boyd, P. W. and Ellwood, M. J. (2010). The biogeochemical cycle of iron in the ocean. *Nature Geoscience*, 3(10):675–682.
- Boyd, P. W., Jickells, T., Law, C. S., Blain, S., Boyle, E. A., Buesseler, K. O., Coale, K. H., Cullen, J. J., de Baar, H. J. W., Follows, M., Harvey, M., Lancelot, C., Levasseur, M., Owens, N. P. J., Pollard, R., Rivkin, R. B., Sarmiento, J., Schoemann, V., Smetacek, V., Takeda, S., Tsuda, A., Turner, S., and Watson, A. J. (2007). Mesoscale Iron Enrichment Experiments 1993-2005: Synthesis and Future Directions. *Science*, 315(5812):612–617.
- Boyle, E., Edmond, J., and Sholkovitz, E. (1977). The mechanism of iron removal in estuaries. *Geochimica et Cosmochimica Acta*, 41(9):1313–1324.

- Bralower, T. J., Meissner, K. J., Alexander, K., and Thomas, D. J. (2014). The dynamics of global change at the Paleocene-Eocene thermal maximum: A data-model comparison. *Geochemistry, Geophysics, Geosystems*, 15(10):3830–3848.
- Bruland, K. W., Franks, R. P., Knauer, G. A., and Martin, J. H. (1979). Sampling and analytical methods for the determination of copper, cadmium, zinc, and nickel at the nanogram per liter level in sea water. *Analytica Chimica Acta*, 105(C):233–245.
- Brussaard, C. P., Noordeloos, A. A., and Riegman, R. (1997). Autolysis kinetics of the marine diatom Ditylum brightwellii (Bacillariophyceae) under nitrogen and phosphorus limitation and starvation. *Journal of Phycology*, 33(6):980–987.
- Bucciarelli, E., Ridame, C., Sunda, W. G., Dimier-Hugueney, C., Cheize, M., and Belviso, S. (2013). Increased intracellular concentrations of DMSP and DMSO in iron-limited oceanic phytoplankton Thalassiosira oceanica and Trichodesmium erythraeum. *Limnology and Oceanography*, 58(5):1667–1679.
- Buchanan, P. J., Chase, Z., Matear, R. J., Phipps, S. J., and Bindoff, N. L. (2019). Marine nitrogen fixers mediate a low latitude pathway for atmospheric CO<sub>2</sub> drawdown. *Nature Communications*, 10(1):1–10.
- Buck, K. N., Lohan, M. C., Sander, S. G., Hassler, C., and Pižeta, I. (2017). Editorial: Organic Ligands—A Key Control on Trace Metal Biogeochemistry in the Ocean. *Frontiers in Marine Science*, 4(September):1–2.
- Buck, K. N., Sohst, B., and Sedwick, P. N. (2015). The organic complexation of dissolved iron along the U.S. GEOTRACES (GA03) North Atlantic Section. *Deep Sea Research Part II: Topical Studies in Oceanography*, 116(September 2016):152–165.
- Buesseler, K. O., Doney, S. C., Karl, D. M., Boyd, P. W., Caldeira, K., Chai, F., Coale, K. H., de Baar, H. J. W., Falkowski, P. G., Johnson, K. S., Lampitt, R. S., Michaels, A. F., Naqvi, S. W. a., Smetacek, V., Takeda, S., and Watson, A. J. (2008). ENVIRON-MENT: Ocean Iron Fertilization–Moving Forward in a Sea of Uncertainty. *Science*, 319(5860):162–162.
- Cabré, A., Marinov, I., and Leung, S. (2015). Consistent global responses of marine ecosystems to future climate change across the IPCC AR5 earth system models. *Climate Dynamics*, 45(5-6):1253–1280.
- Calbet, A. and Landry, M. R. (2004). Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnology and Oceanography*, 49(1):51–57.
- Canfield, D. E., Rosing, M. T., and Bjerrum, C. (2006). Early anaerobic metabolisms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1474):1819–1836.

- Carr, M.-E., Friedrichs, M. A., Schmeltz, M., Noguchi Aita, M., Antoine, D., Arrigo, K. R., Asanuma, I., Aumont, O., Barber, R., Behrenfeld, M., Bidigare, R., Buitenhuis, E. T., Campbell, J., Ciotti, A., Dierssen, H., Dowell, M., Dunne, J., Esaias, W., Gentili, B., Gregg, W., Groom, S., Hoepffner, N., Ishizaka, J., Kameda, T., Le Quéré, C., Lohrenz, S., Marra, J., Mélin, F., Moore, K., Morel, A., Reddy, T. E., Ryan, J., Scardi, M., Smyth, T., Turpie, K., Tilstone, G., Waters, K., and Yamanaka, Y. (2006). A comparison of global estimates of marine primary production from ocean color. *Deep Sea Research Part II: Topical Studies in Oceanography*, 53(5-7):741–770.
- Chen, Y., Tovar-Sanchez, A., Siefert, R. L., Sañudo-Wilhelmy, S. A., and Zhuang, G. (2011). Luxury uptake of aerosol iron by Trichodesmium in the western tropical North Atlantic. *Geophysical Research Letters*, 38(18):n/a–n/a.
- Claussen, M., Mysak, L., Weaver, A., Crucifix, M., Fichefet, T., Loutre, M.-F., Weber, S., Alcamo, J., Alexeev, V., Berger, A., Calov, R., Ganopolski, A., Goosse, H., Lohmann, G., Lunkeit, F., Mokhov, I., Petoukhov, V., Stone, P., and Wang, Z. (2002). Earth system models of intermediate complexity: closing the gap in the spectrum of climate system models. *Climate Dynamics*, 18(7):579–586.
- Conroy, B. J., Steinberg, D. K., Song, B., Kalmbach, A., Carpenter, E. J., and Foster, R. A. (2017). Mesozooplankton graze on cyanobacteria in the Amazon River plume and western tropical North Atlantic. *Frontiers in Microbiology*, 8(AUG):1–15.
- Cullen, J. J. (1995). Status of the iron hypothesis after the Open-Ocean Enrichment Experiment1. *Limnology and Oceanography*, 40(7):1336–1343.
- de Baar, H. J., Boyd, P. W., Coale, K. H., Landry, M. R., Tsuda, A., Assmy, P., Bakker, D. C., Bozec, Y., Barber, R. T., Brzezinski, M. A., Buesseler, K. O., Boyé, M., Croot, P. L., Gervais, F., Gorbunov, M. Y., Harrison, P. J., Hiscock, W. T., Laan, P., Lancelot, C., Law, C. S., Levasseur, M., Marchetti, A., Millero, F. J., Nishioka, J., Nojiri, Y., van Oijen, T., Riebesell, U., Rijkenberg, M. J., Saito, H., Takeda, S., Timmermans, K. R., Veldhuis, M. J., Waite, A. M., and Wong, C. S. (2005). Synthesis of iron fertilization experiments: From the iron age in the age of enlightenment. *Journal of Geophysical Research C: Oceans*, 110(9):1–24.
- Denman, K. (2003). Modelling planktonic ecosystems: parameterizing complexity. *Progress in Oceanography*, 57(3-4):429–452.
- Deutsch, C., Sarmiento, J. L., Sigman, D. M., Gruber, N., and Dunne, J. P. (2007). Spatial coupling of nitrogen inputs and losses in the ocean. *Nature*, 445(7124):163–167.
- Doney, S. C. (1999). Major challenges confronting marine biogeochemical modeling. *Global Biogeochemical Cycles*, 13(3):705–714.
- Doney, S. C., Glover, D. M., and Najjar, R. G. (1996). A new coupled, one-dimensional biological-physical model for the upper ocean: Applications to the JGOFS Bermuda Atlantic Time-series Study (BATS) site. *Deep Sea Research Part II: Topical Studies in Oceanography*, 43(2-3):591–624.

- Ducklow, H. W. (1999). The bacterial component of the oceanic euphotic zone. *FEMS Microbiology Ecology*, 30(1):1–10.
- Dunne, J. P., Sarmiento, J. L., and Gnanadesikan, A. (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):n/a–n/a.
- Eby, M., Weaver, A. J., Alexander, K., Zickfeld, K., Abe-Ouchi, A., Cimatoribus, A. A., Crespin, E., Drijfhout, S. S., Edwards, N. R., Eliseev, A. V., Feulner, G., Fichefet, T., Forest, C. E., Goosse, H., Holden, P. B., Joos, F., Kawamiya, M., Kicklighter, D., Kienert, H., Matsumoto, K., Mokhov, I. I., Monier, E., Olsen, S. M., Pedersen, J. O. P., Perrette, M., Philippon-Berthier, G., Ridgwell, A., Schlosser, A., Schneider von Deimling, T., Shaffer, G., Smith, R. S., Spahni, R., Sokolov, A. P., Steinacher, M., Tachiiri, K., Tokos, K., Yoshimori, M., Zeng, N., and Zhao, F. (2013). Historical and idealized climate model experiments: an intercomparison of Earth system models of intermediate complexity. *Climate of the Past*, 9(3):1111–1140.
- Edwards, K. F., Thomas, M. K., Klausmeier, C. A., and Litchman, E. (2015). Light and growth in marine phytoplankton: Allometric, taxonomic, and environmental variation. *Limnology and Oceanography*, 60(2):540–552.
- Elrod, V. A., Berelson, W. M., Coale, K. H., and Johnson, K. S. (2004). The flux of iron from continental shelf sediments: A missing source for global budgets. *Geophysical Research Letters*, 31(12).
- Eppley, R. W. (1972). Temperature and Phytoplankton Growth in the Sea. *Fishery Bulletin*, 70(4):1063–1085.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958.
- Fanning, A. F. and Weaver, A. J. (1996). An atmospheric energy-moisture balance model: Climatology, interpentadal climate change, and coupling to an ocean general circulation model. *Journal of Geophysical Research*, 101:15111.
- Fasham, M. J. (1995). Variations in the seasonal cycle of biological production in subarctic oceans: A model sensitivity analysis. *Deep-Sea Research Part I*, 42(7):1111–1149.
- Fenchel, T. (1982). Ecology of Heterotrophic Microflagellates. IV Quantitative Occurrence and Importance as Bacterial Consumers. *Marine Ecology Progress Series*, 9(1977):35–42.
- Fitzsimmons, J. N., Carrasco, G. G., Wu, J., Roshan, S., Hatta, M., Measures, C. I., Conway, T. M., John, S. G., and Boyle, E. A. (2015). Partitioning of dissolved iron and iron isotopes into soluble and colloidal phases along the GA03 GEOTRACES

North Atlantic Transect. Deep-Sea Research Part II: Topical Studies in Oceanography, 116:130–151.

- Friedrichs, M. A. M., Dusenberry, J. A., Anderson, L. A., Armstrong, R. A., Chai, F., Christian, J. R., Doney, S. C., Dunne, J., Fujii, M., Hood, R., McGillicuddy, D. J., Moore, J. K., Schartau, M., Spitz, Y. H., and Wiggert, J. D. (2007). Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. *Journal of Geophysical Research*, 112(C8):C08001.
- Frouin, R. and Pinker, R. T. (1995). Estimating Photosynthetically Active Radiation (PAR) at the earth's surface from satellite observations. *Remote Sensing of Environment*, 51(1):98–107.
- Fu, F., Yu, E., Garcia, N., Gale, J., Luo, Y., Webb, E., and Hutchins, D. (2014). Differing responses of marine N<sub>2</sub> fixers to warming and consequences for future diazotroph community structure. *Aquatic Microbial Ecology*, 72(1):33–46.
- Fu, W., Randerson, J. T., and Keith Moore, J. (2016). Climate change impacts on net primary production (NPP) and export production (EP) regulated by increasing stratification and phytoplankton community structure in the CMIP5 models. *Biogeosciences*, 13(18):5151–5170.
- Galbraith, E. D., Gnanadesikan, A., Dunne, J. P., and Hiscock, M. R. (2010). Regional impacts of iron-light colimitation in a global biogeochemical model. *Biogeosciences*, 7(3):1043–1064.
- Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O. K., Zweng, M. M., Reagan, J. R., and Johnson, D. R. (2013a). World Ocean Atlas 2013, Volume 4 : Dissolved Inorganic Nutrients (phosphate, nitrate, silicate). NOAA Atlas NESDIS 76, 4(September):27.
- Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Mishonov, A. V., Baranova, O. K., Zweng, M. M., Reagan, J. R., and Johnson, D. R. (2013b). World Ocean Atlas 2013. Vol. 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. *Technical Ed. NOAA Atlas NESDIS 75*, 3(September):27.
- Garcia, N. S., Fu, F., Sedwick, P. N., and Hutchins, D. A. (2015). Iron deficiency increases growth and nitrogen-fixation rates of phosphorus-deficient marine cyanobacteria. *The ISME Journal*, 9(1):238–245.
- Geider, R. J. and La Roche, J. (2002). Redfield revisited: Variability of C:N:P in marine microalgae and its biochemical basis. *European Journal of Phycology*, 37(1):1–17.
- Geider, R. J., MacIntyre, H. L., and Kana, T. M. (1997). Dynamic model of phytoplankton growth and acclimation: Responses of the balanced growth rate and the chlorophyll a:carbon ratio to light, nutrient-limitation and temperature. *Marine Ecology Progress Series*, 148(1-3):187–200.

- Gledhill, M. (2012). The organic complexation of iron in the marine environment: a review. *Frontiers in Microbiology*, 3(FEB):1–17.
- Gledhill, M. and van den Berg, C. M. (1994). Determination of complexation of iron(III) with natural organic complexing ligands in seawater using cathodic stripping voltammetry. *Marine Chemistry*, 47(1):41–54.
- Griffies, S. M., Harrison, M. J., Pacanowski, R. C., and Rosati, A. (2004). A technical guide to MOM4. *GFDL Ocean Group Tech. Rep*, 5(5):371.
- Grossart, H. P. and Ploug, H. (2001). Microbial degradation of organic carbon and nitrogen on diatom aggregates. *Limnol. Oceanogr.*, 46(2):267–277.
- Gruber, N. and Galloway, J. N. (2008). An Earth-system perspective of the global nitrogen cycle. *Nature*, 451(7176):293–296.
- Hain, M. P., Sigman, D. M., and Haug, G. H. (2010). Carbon dioxide effects of Antarctic stratification, North Atlantic Intermediate Water formation, and subantarctic nutrient drawdown during the last ice age: Diagnosis and synthesis in a geochemical box model. *Global Biogeochemical Cycles*, 24(4):n/a–n/a.
- Hamilton, D. S., Moore, J. K., Arneth, A., Bond, T. C., Carslaw, K. S., Hantson, S., Ito, A., Kaplan, J. O., Lindsay, K., Nieradzik, L., Rathod, S. D., Scanza, R. A., and Mahowald, N. M. (2020). Impact of Changes to the Atmospheric Soluble Iron Deposition Flux on Ocean Biogeochemical Cycles in the Anthropocene. *Global Biogeochemical Cycles*, 34(3):1–22.
- Hansell, D. A. (2013). Recalcitrant Dissolved Organic Carbon Fractions. *Annual Review* of Marine Science, 5(1):421–445.
- Hansen, N. (2006). The CMA Evolution Strategy: A Comparing Review. In Jose A., L., Pedro, L., Iñaki, I., and Endika, B., editors, *Towards a New Evolutionary Computation*, volume 192, pages 75–102. Springer-Verlag, Berlin/Heidelberg.
- Hansen, N. and Ostermeier, A. (2001). {C}ompletely {D}erandomized {S}elf-{A}daptation in {E}volution {S}trategies. *Evolutionary Computation*, 9(2):159–195.
- Hawser, S. P., O'Neil, J. M., Roman, M. R., and Codd, G. A. (1992). Toxicity of blooms of the cyanobacterium Trichodesmium to zooplankton. *Journal of Applied Phycology*, 4(1):79–86.
- Hirst, A. G. and Kiørboe, T. (2002). Mortality of marine planktonic copepods: Global rates and patterns. *Marine Ecology Progress Series*, 230:195–209.
- Holland, W. R., Chow, J. C., and Bryan, F. O. (1998). Application of a Third-Order Upwind Scheme in the NCAR Ocean Model\*. *Journal of Climate*, 11(6):1487–1493.

- Honeyman, B. D., Balistrieri, L. S., and Murray, J. W. (1988). Oceanic trace metal scavenging: the importance of particle concentration. *Deep Sea Research Part A. Oceanographic Research Papers*, 35(2):227–246.
- Honjo, S., Manganini, S. J., Krishfield, R. A., and Francois, R. (2008). Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: A synthesis of global sediment trap programs since 1983. *Progress in Oceanography*, 76(3):217–285.
- Hopkinson, B. M. and Barbeau, K. A. (2007). Organic and redox speciation of iron in the eastern tropical North Pacific suboxic zone. *Marine Chemistry*, 106(1-2 SPEC. ISS.):2–17.
- Hopwood, M. J., Carroll, D., Höfer, J., Achterberg, E. P., Meire, L., Le Moigne, F. A. C., Bach, L. T., Eich, C., Sutherland, D. A., and González, H. E. (2019). Highly variable iron content modulates iceberg-ocean fertilisation and potential carbon export. *Nature Communications*, 10(1):5261.
- Iwade, S., Kuma, K., Isoda, Y., Yoshida, M., Kudo, I., Nishioka, J., and Suzuki, K. (2006). Effect of high iron concentrations on iron uptake and growth of a coastal diatom Chaetoceros sociale. *Aquatic Microbial Ecology*, 43(2):177–191.
- Jabre, L. and Bertrand, E. M. (2020). Interactive effects of iron and temperature on the growth of Fragilariopsis cylindrus . *Limnology and Oceanography Letters*, (Eppley 1972).
- Jackson, G. A., Checkley, D. M., and Dagg, M. (2015). Settling of particles in the upper 100 m of the ocean detected with autonomous profiling floats off California. *Deep Sea Research Part I: Oceanographic Research Papers*, 99:75–86.
- Jacq, V., Ridame, C., L'Helguen, S., Kaczmar, F., and Saliot, A. (2014). Response of the Unicellular Diazotrophic Cyanobacterium Crocosphaera watsonii to Iron Limitation. *PLoS ONE*, 9(1):e86749.
- Jin., X., Gruber, N., Dune, J. P., Sarmiento, J. L., and Armstrong, R. A. (2006). Diagnosing the contributions of phytoplankton functional groups to the production and export of particulate organic carbon, CaCO<sub>3</sub>, and opal from global nutrient and alkalinity distributions. *Global Biogeochemical Cycles*, 20(2):1–17.
- Joos, F., Sarmiento, J. L., and Siegenthaler, U. (1991). Estimates of the effect of Southern Ocean iron fertilization on atmospheric CO<sub>2</sub> concentrations. *Nature*, 349(6312):772–775.
- Karl, D., Michaels, A., Bergman, B., Capone, D., Carpenter, E., Letelier, R., Lipschultz, F., Paerl, H., Sigman, D., and Stal, L. (2002). Dinitrogen fixation in the world's oceans. In *Biogeochemistry*.

- Keller, D. P., Oschlies, A., and Eby, M. (2012). A new marine ecosystem model for the University of Victoria Earth System Climate Model. *Geoscientific Model Development*, 5(5):1195–1220.
- Khatiwala, S. (2007). A computational framework for simulation of biogeochemical tracers in the ocean. *Global Biogeochemical Cycles*, 21(3):n/a–n/a.
- Kleppel, G. S. (1993). On the diets of calanoid copepods. *Marine Ecology Progress* Series, 99(1-2):183–195.
- Knapp, A., Dekaezemacker, J., Bonnet, S., Sohm, J., and Capone, D. (2012). Sensitivity of Trichodesmium erythraeum and Crocosphaera watsonii abundance and N<sub>2</sub> fixation rates to varying  $NO_3^-$  and  $PO_4^{3-}$  concentrations in batch cultures. *Aquatic Microbial Ecology*, 66(3):223–236.
- Koeve, W. and Kähler, P. (2010). Heterotrophic denitrification vs. autotrophic anammox quantifying collateral effects on the oceanic carbon cycle. *Biogeosciences*, 7(8):2327–2337.
- Körtzinger, A., Hedges, J. I., and Quay, P. D. (2001). Redfield ratios revisited: Removing the biasing effect of anthropogenic CO<sub>2</sub>. *Limnology and Oceanography*, 46(4):964–970.
- Kriest, I. (2017). Calibration of a simple and a complex model of global marine biogeochemistry. *Biogeosciences Discussions*, pages 1–28.
- Kriest, I., Kähler, P., Koeve, W., Kvale, K., Sauerland, V., and Oschlies, A. (2020). One size fits all? Calibrating an ocean biogeochemistry model for different circulations. *Biogeosciences*, 17(12):3057–3082.
- Kriest, I. and Oschlies, A. (2015). MOPS-1.0: towards a model for the regulation of the global oceanic nitrogen budget by marine biogeochemical processes. *Geoscientific Model Development*, 8(9):2929–2957.
- Kriest, I., Oschlies, A., and Khatiwala, S. (2012). Sensitivity analysis of simple global marine biogeochemical models. *Global Biogeochemical Cycles*, 26(2):n/a–n/a.
- Kriest, I., Sauerland, V., Khatiwala, S., Srivastav, A., and Oschlies, A. (2017). Calibrating a global three-dimensional biogeochemical ocean model (MOPS-1.0). *Geoscientific Model Development*, 10(1):127–154.
- Kustka, A., Saudo-Wilhelmy, S., Carpenter, E. J., Capone, D. G., and Raven, J. A. (2003a). A revised estimate of the iron use efficiency of nitrogen fixation, with special reference to the marine Cyanobacterium Trichodesmium SPP. (Cyanophyta) 1. *Journal of Phycology*, 39(1):12–25.

- Kustka, A. B., Sañudo-Wilhelmy, S. A., Carpenter, E. J., Capone, D., Burns, J., and Sunda, W. G. (2003b). Iron requirements for dinitrogen- And ammonium-supported growth in cultures of Trichodesmium (IMS 101): Comparison with nitrogen fixation rates and iron: Carbon ratios of field populations. *Limnology and Oceanography*, 48(5):1869–1884.
- Kvale, K. F., Khatiwala, S., Dietze, H., Kriest, I., and Oschlies, A. (2017). Evaluation of the transport matrix method for simulation of ocean biogeochemical tracers. *Geoscientific Model Development*, 10(6):2425–2445.
- Kvale, K. F. and Meissner, K. J. (2017). Primary production sensitivity to phytoplankton light attenuation parameter increases with transient forcing. *Biogeosciences*, 14(20):4767–4780.
- Kvale, K. F., Meissner, K. J., and Keller, D. P. (2015). Potential increasing dominance of heterotrophy in the global ocean. *Environmental Research Letters*, 10(7):074009.
- Kvale, K. F., Turner, K. E., Keller, D. P., and Meissner, K. J. (2018). Asymmetric dynamical ocean responses in warming icehouse and cooling greenhouse climates. *Environmental Research Letters*, 13(12):125011.
- Kvale, K. F., Turner, K. E., Landolfi, A., and Meissner, K. J. (2019). Phytoplankton calcifiers control nitrate cycling and the pace of transition in warming icehouse and cooling greenhouse climates. *Biogeosciences*, 16(5):1019–1034.
- Kwiatkowski, L., Yool, A., Allen, J. I., Anderson, T. R., Barciela, R., Buitenhuis, E. T., Butenschön, M., Enright, C., Halloran, P. R., Le Quéré, C., de Mora, L., Racault, M.-F., Sinha, B., Totterdell, I. J., and Cox, P. M. (2014). iMarNet: an ocean biogeochemistry model intercomparison project within a common physical ocean modelling framework. *Biogeosciences*, 11(24):7291–7304.
- Lambert, F., Tagliabue, A., Shaffer, G., Lamy, F., Winckler, G., Farias, L., Gallardo, L., and De Pol-Holz, R. (2015). Dust fluxes and iron fertilization in Holocene and Last Glacial Maximum climates. *Geophysical Research Letters*, 42(14):6014–6023.
- Landolfi, A., Dietze, H., Koeve, W., and Oschlies, A. (2013). Overlooked runaway feedback in the marine nitrogen cycle: the vicious cycle. *Biogeosciences*, 10(3):1351–1363.
- Landolfi, A., Kähler, P., Koeve, W., and Oschlies, A. (2018). Global Marine N<sub>2</sub> Fixation Estimates: From Observations to Models. *Frontiers in Microbiology*, 9(SEP):1–8.
- Landolfi, A., Somes, C. J., Koeve, W., Zamora, L. M., and Oschlies, A. (2017). Oceanic nitrogen cycling and N<sub>2</sub>O flux perturbations in the Anthropocene. *Global Biogeochemical Cycles*, 31(8):1236–1255.
- Laufkötter, C., Vogt, M., Gruber, N., Aita-Noguchi, M., Aumont, O., Bopp, L., Buitenhuis, E., Doney, S. C., Dunne, J., Hashioka, T., Hauck, J., Hirata, T., John, J., Le Quéré, C., Lima, I. D., Nakano, H., Seferian, R., Totterdell, I., Vichi, M., and Völker, C.

(2015). Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences*, 12(23):6955–6984.

- Laufkötter, C., Vogt, M., Gruber, N., Aumont, O., Bopp, L., Doney, S. C., Dunne, J. P., Hauck, J., John, J. G., Lima, I. D., Seferian, R., and Völker, C. (2016). Projected decreases in future marine export production: the role of the carbon flux through the upper ocean ecosystem. *Biogeosciences*, 13(13):4023–4047.
- Le Quéré, C., Harrison, S. P., Colin Prentice, I., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Cotrim Da Cunha, L., Geider, R., Giraud, X., Klaas, C., Kohfeld, K. E., Legendre, L., Manizza, M., Platt, T., Rivkin, R. B., Sathyendranath, S., Uitz, J., Watson, A. J., and Wolf-Gladrow, D. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, 11(11):2016–2040.
- Lefèvre, N. and Watson, A. J. (1999). Modeling the geochemical cycle of iron in the oceans and its impact on atmospheric CO<sub>2</sub> concentrations. *Global Biogeochemical Cycles*, 13(3):727–736.
- Lerner, P., Marchal, O., Lam, P. J., Anderson, R. F., Buesseler, K., Charette, M. A., Edwards, R. L., Hayes, C. T., Huang, K.-F., Lu, Y., Robinson, L. F., and Solow, A. (2016). Corrigendum to "Testing models of thorium and particle cycling in the ocean using data from station GT11-22 of the U.S. GEOTRACES North Atlantic Section" [Deep-Sea Res. I 113 (2016) 57–79]. *Deep Sea Research Part I: Oceanographic Research Papers*, 118:101.
- Litchman, E., Klausmeier, C. A., Schofield, O. M., and Falkowski, P. G. (2007). The role of functional traits and trade-offs in structuring phytoplankton communities: Scaling from cellular to ecosystem level. *Ecology Letters*, 10(12):1170–1181.
- Löptien, U. and Dietze, H. (2015). Constraining parameters in marine pelagic ecosystem models - Is it actually feasible with typical observations of standing stocks? *Ocean Science*, 11(4):573–590.
- Löptien, U. and Dietze, H. (2017). Effects of parameter indeterminacy in pelagic biogeochemical modules of Earth System Models on projections into a warming future: The scale of the problem. *Global Biogeochemical Cycles*, 31(7):1155–1172.
- Louropoulou, E., Gledhill, M., Browning, T. J., Desai, D. K., Barraqueta, J. L. M., Tonnard, M., Sarthou, G., Planquette, H., Bowie, A. R., Schmitz, R. A., LaRoche, J., and Achterberg, E. P. (2019). Regulation of the phytoplankton heme B iron pool during the North Atlantic spring bloom. *Frontiers in Microbiology*, 10(JULY).
- Luo, C., Mahowald, N., Bond, T., Chuang, P. Y., Artaxo, P., Siefert, R., Chen, Y., and Schauer, J. (2008). Combustion iron distribution and deposition. *Global Biogeochemical Cycles*, 22(1):1–17.

- Lutz, M. J., Caldeira, K., Dunbar, R. B., and Behrenfeld, M. J. (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*, 112(C10):C10011.
- Martin, J. H. (1990). Glacial-interglacial CO<sub>2</sub> change: The Iron Hypothesis. *Paleo-ceanography*, 5(1):1–13.
- Martin, J. H., Knauer, G. A., Karl, D. M., and Broenkow, W. W. (1987). VERTEX: carbon cycling in the northeast Pacific. *Deep Sea Research Part A, Oceanographic Research Papers*, 34(2):267–285.
- Martiny, A. C., Vrugt, J. A., Primeau, F. W., and Lomas, M. W. (2013). Regional variation in the particulate organic carbon to nitrogen ratio in the surface ocean. *Global Biogeochemical Cycles*, 27(3):723–731.
- Meissner, K. J., Galbraith, E. D., and Völker, C. (2005). Denitrification under glacial and interglacial conditions: A physical approach. *Paleoceanography*, 20(3):n/a–n/a.
- Meissner, K. J., Weaver, A. J., Matthews, H. D., and Cox, P. M. (2003). The role of land surface dynamics in glacial inception: A study with the UVic Earth System Model. *Climate Dynamics*, 21(7-8):515–537.
- Mengis, N., Keller, D. P., Rickels, W., Quaas, M., and Oschlies, A. (2019). Climate engineering-induced changes in correlations between Earth system variables—implications for appropriate indicator selection. *Climatic Change*, 153(3):305–322.
- Mills, M. M., Turk-Kubo, K. A., van Dijken, G. L., Henke, B. A., Harding, K., Wilson, S. T., Arrigo, K. R., and Zehr, J. P. (2020). Unusual marine cyanobacteria/haptophyte symbiosis relies on N<sub>2</sub> fixation even in N-rich environments. *The ISME Journal*, 14(10):2395–2406.
- Moisander, P. H., Benavides, M., Bonnet, S., Berman-Frank, I., White, A. E., and Riemann, L. (2017). Chasing after Non-cyanobacterial Nitrogen Fixation in Marine Pelagic Environments. *Frontiers in Microbiology*, 8(SEP).
- Moore, C. M., Mills, M. M., Achterberg, E. P., Geider, R. J., LaRoche, J., Lucas, M. I., McDonagh, E. L., Pan, X., Poulton, A. J., Rijkenberg, M. J. A., Suggett, D. J., Ussher, S. J., and Woodward, E. M. S. (2009). Large-scale distribution of Atlantic nitrogen fixation controlled by iron availability. *Nature Geoscience*, 2(12):867–871.
- Moore, C. M., Mills, M. M., Arrigo, K. R., Berman-Frank, I., Bopp, L., Boyd, P. W., Galbraith, E. D., Geider, R. J., Guieu, C., Jaccard, S. L., Jickells, T. D., La Roche, J., Lenton, T. M., Mahowald, N. M., Marañón, E., Marinov, I., Moore, J. K., Nakatsuka, T., Oschlies, A., Saito, M. A., Thingstad, T. F., Tsuda, A., and Ulloa, O. (2013). Processes and patterns of oceanic nutrient limitation. *Nature Geoscience*, 6(9):701–710.

- Moore, J. K., Braucher, O., and Sedimentary, O. B. (2008). Sedimentary and mineral dust sources of dissolved iron to the world ocean To cite this version : Biogeosciences Sedimentary and mineral dust sources of dissolved iron to the world ocean.
- Moore, J. K., Doney, S. C., and Lindsay, K. (2004). Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. *Global Biogeochemical Cycles*, 18(4):1–21.
- Mulholland, M. R. (2007). The fate of nitrogen fixed by diazotrophs in the ocean. *Bio-geosciences*, 4(1):37–51.
- Nickelsen, L., Keller, D. P., and Oschlies, A. (2015). A dynamic marine iron cycle module coupled to the University of Victoria Earth System Model: the Kiel Marine Biogeochemical Model 2 for UVic 2.9. *Geoscientific Model Development*, 8(5):1357–1381.
- Oschlies, A. and Garçon, V. (1999). An Eddy-Permitting Coupled Physical-Biological Model of the North Atlantic: 1. Sensitivity to Advection Numerics and Mixed Layer Physics. *Global Biogeochem. CyclesGlobal Bi*, 13(1):135–160.
- Oschlies, A., Koeve, W., Landolfi, A., and Kähler, P. (2019). Loss of fixed nitrogen causes net oxygen gain in a warmer future ocean. *Nature Communications*.
- Oschlies, A. and Schartau, M. (2005). Basin-scale performance of a locally optimized marine ecosystem model. *Journal of Marine Research*, 63(2):335–358.
- Pacanowski, R. C. (1996). MOM2. Documentation, user's guide and reference manual. *GFDL Ocean Group Technical Report 3.2.*
- Parekh, P., Follows, M. J., and Boyle, E. (2004). Modeling the global ocean iron cycle. *Global Biogeochemical Cycles*, 18(1):n/a–n/a.
- Parekh, P., Follows, M. J., and Boyle, E. A. (2005). Decoupling of iron and phosphate in the global ocean. *Global Biogeochemical Cycles*, 19(2):1–16.
- Parsons, T. R., Takahashi, M., and Hargrave, B. (1984). Biological oceanographic processes. 3rd edition. *Biological oceanographic processes. 3rd edition.*
- Pasquier, B. and Holzer, M. (2017). Inverse-model estimates of the ocean's coupled phosphorus, silicon, and iron cycles. *Biogeosciences*, 14(18):4125–4159.
- Passow, U. and Carlson, C. (2012). The biological pump in a high CO<sub>2</sub> world. *Marine Ecology Progress Series*, 470(2):249–271.
- Pauly, D. and Christensen, V. (1995). Primary production required to sustain global fisheries. *Nature*, 374(6519):255–257.
- Pelegrí, S. P., Christaki, U., Dolan, J., and Rassoulzadegan, F. (1999). Particulate and dissolved organic carbon production by the heterotrophic nanoflagellate Pteridomonas danica Patterson and Fenchel. *Microbial Ecology*, 37(4):276–284.

- Pham, A. L. D. and Ito, T. (2018). Formation and Maintenance of the GEOTRACES Subsurface-Dissolved Iron Maxima in an Ocean Biogeochemistry Model. *Global Bio-geochemical Cycles*, 32(6):932–953.
- Pham, A. L. D. and Ito, T. (2019). Ligand Binding Strength Explains the Distribution of Iron in the North Atlantic Ocean. *Geophysical Research Letters*, 46(13):7500–7508.
- Pitchford, J. (1999). Iron limitation, grazing pressure and oceanic high nutrient-low chlorophyll (HNLC) regions. *Journal of Plankton Research*, 21(3):525–547.
- Pomeroy, L. R. (2001). Caught in the food web: Complexity made simple? *Scientia Marina*, 65(SUPPLEMENT 2):31–40.
- Pope, R. M. and Fry, E. S. (1997). Absorption spectrum (380-700nm) of pure water. II . Integrating cavity measurements. *Applied Optics*, 36(33):8710–8723.
- Price, N. M. (2005). The elemental stoichiometry and composition of an iron-limited diatom. *Limnology and Oceanography*, 50(4):1159–1171.
- Price, N. M., Ahner, B. A., and Morel, F. M. M. (1994). The equatorial Pacific Ocean: Grazer-controlled phytoplankton populations in an iron-limited ecosystem1. *Limnology and Oceanography*, 39(3):520–534.
- Primeau, F. W., Holzer, M., and DeVries, T. (2013). Southern Ocean nutrient trapping and the efficiency of the biological pump. *Journal of Geophysical Research: Oceans*, 118(5):2547–2564.
- Raiswell, R., Benning, L. G., Tranter, M., and Tulaczyk, S. (2008). Bioavailable iron in the Southern Ocean: the significance of the iceberg conveyor belt. *Geochemical Transactions*, 9(1):7.
- Raven, J. A. (1988). The iron and molybdenum use efficiencies of plant growth with different energy, carbon and nitrogen sources. *New Phytologist*, 109(3):279–287.
- Redfield, A. C. (1958). The biological control of chemical factors in the environment. *American Scientist*, 46(3):230A–221.
- Reith, F., Keller, D. P., and Oschlies, A. (2016). Revisiting ocean carbon sequestration by direct injection: A global carbon budget perspective. *Earth System Dynamics*, 7(4):797–812.
- Riche, O. G. and Christian, J. R. (2018). Ocean dinitrogen fixation and its potential effects on ocean primary production in Earth system model simulations of anthropogenic warming. *Elementa*, 6.
- Ridgwell, A., Hargreaves, J. C., Edwards, N. R., Annan, J. D., Lenton, T. M., Marsh, R., Yool, A., and Watson, A. (2007). Marine geochemical data assimilation in an efficient Earth System Model of global biogeochemical cycling. *Biogeosciences*, 4(1):87–104.

- Riebesell, U., Kortzinger, A., and Oschlies, A. (2009). Sensitivities of marine carbon fluxes to ocean change. *Proceedings of the National Academy of Sciences*, 106(49):20602–20609.
- Robinson, J., Popova, E. E., Srokosz, M. A., and Yool, A. (2016). A tale of three islands: Downstream natural iron fertilization in the Southern Ocean. *Journal of Geophysical Research: Oceans*, 121(5):3350–3371.
- Rue, E. L. and Bruland, K. W. (1995). Complexation of iron(III) by natural organic ligands in the Central North Pacific as determined by a new competitive ligand equilibration/adsorptive cathodic stripping voltammetric method. *Marine Chemistry*, 50(1-4):117–138.
- Ryther, J. H. and Dunstan, W. M. (1971). Nitrogen, Phosphorus, and Eutrophication in the Coastal Marine Environment. *Science*, 171(3975):1008–1013.
- Sarmiento, J. L., Dunne, J., Gnanadesikan, A., Key, R. M., Matsumoto, K., and Slater, R. (2002). A new estimate of the CaCO<sub>3</sub> to organic carbon export ratio. *Global Biogeochemical Cycles*, 16(4):54–1.
- Sarmiento, J. L. and Gruber, N. (2006). *Ocean Biogeochemical Dynamics*. Princeton University Press, Princeton and Oxford.
- Schartau, M. and Oschlies, A. (2003). Simultaneous data-based optimization of a 1Decosystem model at three locations in the North Atlantic: Part II - Standing stocks and nitrogen fluxes. *Journal of Marine Research*, 61(6):795–821.
- Schlitzer, R., Anderson, R. F., Dodas, E. M., Lohan, M., Geibert, W., Tagliabue, A., Bowie, A., Jeandel, C., Maldonado, M. T., Landing, W. M., Cockwell, D., Abadie, C., Abouchami, W., Achterberg, E. P., Agather, A., Aguliar-Islas, A., van Aken, H. M., Andersen, M., Archer, C., Auro, M., de Baar, H. J., Baars, O., Baker, A. R., Bakker, K., Basak, C., Baskaran, M., Bates, N. R., Bauch, D., van Beek, P., Behrens, M. K., Black, E., Bluhm, K., Bopp, L., Bouman, H., Bowman, K., Bown, J., Boyd, P., Boye, M., Boyle, E. A., Branellec, P., Bridgestock, L., Brissebrat, G., Browning, T., Bruland, K. W., Brumsack, H. J., Brzezinski, M., Buck, C. S., Buck, K. N., Buesseler, K., Bull, A., Butler, E., Cai, P., Mor, P. C., Cardinal, D., Carlson, C., Carrasco, G., Casacuberta, N., Casciotti, K. L., Castrillejo, M., Chamizo, E., Chance, R., Charette, M. A., Chaves, J. E., Cheng, H., Chever, F., Christl, M., Church, T. M., Closset, I., Colman, A., Conway, T. M., Cossa, D., Croot, P., Cullen, J. T., Cutter, G. A., Daniels, C., Dehairs, F., Deng, F., Dieu, H. T., Duggan, B., Dulaquais, G., Dumousseaud, C., Echegoyen-Sanz, Y., Edwards, R. L., Ellwood, M., Fahrbach, E., Fitzsimmons, J. N., Russell Flegal, A., Fleisher, M. Q., van de Flierdt, T., Frank, M., Friedrich, J., Fripiat, F., Fröllje, H., Galer, S. J., Gamo, T., Ganeshram, R. S., Garcia-Orellana, J., Garcia-Solsona, E., Gault-Ringold, M., George, E., Gerringa, L. J., Gilbert, M., Godoy, J. M., Goldstein, S. L., Gonzalez, S. R., Grissom, K., Hammerschmidt, C., Hartman, A., Hassler, C. S., Hathorne, E. C., Hatta, M., Hawco, N., Hayes, C. T., Heimbürger, L. E., Helgoe, J.,

Heller, M., Henderson, G. M., Henderson, P. B., van Heuven, S., Ho, P., Horner, T. J., Hsieh, Y. T., Huang, K. F., Humphreys, M. P., Isshiki, K., Jacquot, J. E., Janssen, D. J., Jenkins, W. J., John, S., Jones, E. M., Jones, J. L., Kadko, D. C., Kayser, R., Kenna, T. C., Khondoker, R., Kim, T., Kipp, L., Klar, J. K., Klunder, M., Kretschmer, S., Kumamoto, Y., Laan, P., Labatut, M., Lacan, F., Lam, P. J., Lambelet, M., Lamborg, C. H., Le Moigne, F. A., Le Roy, E., Lechtenfeld, O. J., Lee, J. M., Lherminier, P., Little, S., López-Lora, M., Lu, Y., Masque, P., Mawji, E., Mcclain, C. R., Measures, C., Mehic, S., Barraqueta, J. L. M., van der Merwe, P., Middag, R., Mieruch, S., Milne, A., Minami, T., Moffett, J. W., Moncoiffe, G., Moore, W. S., Morris, P. J., Morton, P. L., Nakaguchi, Y., Nakayama, N., Niedermiller, J., Nishioka, J., Nishiuchi, A., Noble, A., Obata, H., Ober, S., Ohnemus, D. C., van Ooijen, J., O'Sullivan, J., Owens, S., Pahnke, K., Paul, M., Pavia, F., Pena, L. D., Peters, B., Planchon, F., Planquette, H., Pradoux, C., Puigcorbé, V., Quay, P., Queroue, F., Radic, A., Rauschenberg, S., Rehkämper, M., Rember, R., Remenyi, T., Resing, J. A., Rickli, J., Rigaud, S., Rijkenberg, M. J., Rintoul, S., Robinson, L. F., Roca-Martí, M., Rodellas, V., Roeske, T., Rolison, J. M., Rosenberg, M., Roshan, S., Rutgers van der Loeff, M. M., Ryabenko, E., Saito, M. A., Salt, L. A., Sanial, V., Sarthou, G., Schallenberg, C., Schauer, U., Scher, H., Schlosser, C., Schnetger, B., Scott, P., Sedwick, P. N., Semiletov, I., Shelley, R., Sherrell, R. M., Shiller, A. M., Sigman, D. M., Singh, S. K., Slagter, H. A., Slater, E., Smethie, W. M., Snaith, H., Sohrin, Y., Sohst, B., Sonke, J. E., Speich, S., Steinfeldt, R., Stewart, G., Stichel, T., Stirling, C. H., Stutsman, J., Swarr, G. J., Swift, J. H., Thomas, A., Thorne, K., Till, C. P., Till, R., Townsend, A. T., Townsend, E., Tuerena, R., Twining, B. S., Vance, D., Velazquez, S., Venchiarutti, C., Villa-Alfageme, M., Vivancos, S. M., Voelker, A. H., Wake, B., Warner, M. J., Watson, R., van Weerlee, E., Alexandra Weigand, M., Weinstein, Y., Weiss, D., Wisotzki, A., Woodward, E. M. S., Wu, J., Wu, Y., Wuttig, K., Wyatt, N., Xiang, Y., Xie, R. C., Xue, Z., Yoshikawa, H., Zhang, J., Zhang, P., Zhao, Y., Zheng, L., Zheng, X. Y., Zieringer, M., Zimmer, L. A., Ziveri, P., Zunino, P., and Zurbrick, C. (2018). The GEOTRACES Intermediate Data Product 2017. Chemical Geology, 493(June):210–223.

- Schlosser, C., Klar, J. K., Wake, B. D., Snow, J. T., Honey, D. J., Woodward, E. M. S., Lohan, M. C., Achterberg, E. P., and Moore, C. M. (2014). Seasonal ITCZ migration dynamically controls the location of the (sub)tropical Atlantic biogeochemical divide. *Proceedings of the National Academy of Sciences*, 111(4):1438–1442.
- Schmittner, A., Oschlies, A., Giraud, X., Eby, M., and Simmons, H. L. (2005). A global model of the marine ecosystem for long-term simulations: Sensitivity to ocean mixing, buoyancy forcing, particle sinking, and dissolved organic matter cycling. *Global Biogeochemical Cycles*, 19(3):1–17.
- Schmittner, A., Oschlies, A., Matthews, H. D., and Galbraith, E. D. (2008). Future changes in climate, ocean circulation, ecosystems, and biogeochemical cycling simulated for a business-as-usual CO<sub>2</sub>-emission scenario until year 4000 AD. *Global Biogeochemical Cycles*, 22(1):1–21.

- Scott, V., Kettle, H., and Merchant, C. J. (2011). Sensitivity analysis of an ocean carbon cycle model in the North Atlantic: An investigation of parameters affecting the air-sea CO<sub>2</sub> flux, primary production and export of detritus. *Ocean Science*, 7(3):405–419.
- Sellner, K. G. (1997). Physiology, ecology, and toxic properties of marine cyanobacteria blooms. *Limnology and Oceanography*, 42(5\_part\_2):1089–1104.
- Settle, D. and Patterson, C. (1980). Lead in albacore: guide to lead pollution in Americans. *Science*, 207(4436):1167–1176.
- Shaffer, G. (2010). Long-term effectiveness and consequences of carbon dioxide sequestration. *Nature Geoscience*, 3(7):464–467.
- Shirasawa, K., Ehn, J., Granskog, M., Ishikawa, N., Kawamura, T., and Ice, S. (2001). Ice and Snow Cover As a Filter of Solar Radiation in. In Levin, I. and Gilbert, G., editors, *Current Problems in Optics of Natural Waters*, pages 278–282. DS Rozhdestvensky Opt. Soc., St. Petersburg, Russia.
- Shire, D. M. and Kustka, A. B. (2015). Luxury uptake, iron storage and ferritin abundance in Prochlorococcus marinus (Synechococcales) strain MED4. *Phycologia*, 54(4):398– 406.
- Singh, A., Baer, S. E., Riebesell, U., Martiny, A. C., and Lomas, M. W. (2015). C : N : P stoichiometry at the Bermuda Atlantic Time-series Study station in the North Atlantic Ocean. *Biogeosciences*, 12(21):6389–6403.
- Snow, J. T., Schlosser, C., Woodward, E. M. S., Mills, M. M., Achterberg, E. P., Mahaffey, C., Bibby, T. S., and Moore, C. M. (2015). Environmental controls on the biogeography of diazotrophy and Trichodesmium in the Atlantic Ocean. *Global Biogeochemical Cycles*, 29(6):865–884.
- Stock, C. A., John, J. G., Rykaczewski, R. R., Asch, R. G., Cheung, W. W. L., Dunne, J. P., Friedland, K. D., Lam, V. W. Y., Sarmiento, J. L., and Watson, R. A. (2017). Reconciling fisheries catch and ocean productivity. *Proceedings of the National Academy* of Sciences, 114(8):E1441–E1449.
- Tagliabue, A., Aumont, O., and Bopp, L. (2014). The impact of different external sources of iron on the global carbon cycle. *Geophysical Research Letters*, 41(3):10.1002/2013GL059059.
- Tagliabue, A., Aumont, O., DeAth, R., Dunne, J. P., Dutkiewicz, S., Galbraith, E., Misumi, K., Moore, J. K., Ridgwell, A., Sherman, E., Stock, C., Vichi, M., Völker, C., and Yool, A. (2016). How well do global ocean biogeochemistry models simulate dissolved iron distributions? *Global Biogeochemical Cycles*, 30(2):149–174.
- Tagliabue, A., Bopp, L., Dutay, J.-C., Bowie, A. R., Chever, F., Jean-Baptiste, P., Bucciarelli, E., Lannuzel, D., Remenyi, T., Sarthou, G., Aumont, O., Gehlen, M., and Jeandel,

C. (2010). Hydrothermal contribution to the oceanic dissolved iron inventory. *Nature Geoscience*, 3(4):252–256.

- Tagliabue, A., Bowie, A. R., Boyd, P. W., Buck, K. N., Johnson, K. S., and Saito, M. A. (2017). The integral role of iron in ocean biogeochemistry. *Nature*, 543(7643):51–59.
- Tagliabue, A., Mtshali, T., Aumont, O., Bowie, A. R., Klunder, M. B., Roychoudhury, A. N., and Swart, S. (2012). A global compilation of dissolved iron measurements: Focus on distributions and processes in the Southern Ocean. *Biogeosciences*, 9(6):2333– 2349.
- Taucher, J., Bach, L. T., Riebesell, U., and Oschlies, A. (2014). The viscosity effect on marine particle flux: A climate relevant feedback mechanism. *Global Biogeochemical Cycles*, 28(4):415–422.
- Taucher, J. and Oschlies, A. (2011). Can we predict the direction of marine primary production change under global warming? *Geophysical Research Letters*, 38(2):1–6.
- Taylor, K. E. (2001). Summarizing Multiple Aspects of Model Performance in a Single Diagram. *Journal of Geophysical Research*, 106(D7):7183–7192.
- Timmermans, K. R., van der Wagt, B., and de Baar, H. J. W. (2004). Growth rates, halfsaturation constants, and silicate, nitrate, and phosphate depletion in relation to iron availability of four large, open-ocean diatoms from the Southern Ocean. *Limnology* and Oceanography, 49(6):2141–2151.
- Twining, B. S., Antipova, O., Chappell, P. D., Cohen, N. R., Jacquot, J. E., Mann, E. L., Marchetti, A., Ohnemus, D. C., Rauschenberg, S., and Tagliabue, A. (2021). Taxonomic and nutrient controls on phytoplankton iron quotas in the ocean. *Limnology and Oceanography Letters*, 6(2):96–106.
- Twining, B. S. and Baines, S. B. (2013). The Trace Metal Composition of Marine Phytoplankton. *Annual Review of Marine Science*, 5(1):191–215.
- Tyrrell, T. (1999). The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature*, 400(6744):525–531.
- Tyrrell, T. and Law, C. S. (1997). Low nitrate:phosphate ratios in the global ocean. *Nature*, 387(6635):793–796.
- van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K. (2011). The representative concentration pathways: an overview. *Climatic Change*, 109(1-2):5–31.
- Volk, T. and Hoffert, M. I. (1985). Ocean carbon pumps: analysis of relative strengths and efficiencies in ocean-driven atmospheric CO<sub>2</sub> changes. In Sundquist, E. T. and Broecker, W. S., editors, *The carbon cycle and atmospheric CO<sub>2</sub> : natural variations*

Archean to present. Chapman conference papers, 1984, pages 99–110. American Geophysical Union; Geophysical Monograph 32.

- Völker, C. and Tagliabue, A. (2015). Modeling organic iron-binding ligands in a threedimensional biogeochemical ocean model. *Marine Chemistry*, 173:67–77.
- Wang, W. L., Moore, J. K., Martiny, A. C., and Primeau, F. W. (2019). Convergent estimates of marine nitrogen fixation. *Nature*, 566(7743):205–211.
- Ward, B. A., Dutkiewicz, S., Moore, C. M., and Follows, M. J. (2013). Iron, phosphorus, and nitrogen supply ratios define the biogeography of nitrogen fixation. *Limnology and Oceanography*, 58(6):2059–2075.
- Watson, A. J., Bakker, D. C. E., Ridgwell, A. J., Boyd, P. W., and Law, C. S. (2000). Effect of iron supply on Southern Ocean CO<sub>2</sub> uptake and implications for glacial atmospheric CO<sub>2</sub>. *Nature*, 407(6805):730–733.
- Weaver, A. J. and Eby, M. (1997). On the Numerical Implementation of Advection Schemes for Use in Conjunction with Various Mixing Parameterizations in the GFDL Ocean Model. *Journal of Physical Oceanography*, 27(2):369–377.
- Weaver, A. J., Eby, M., Wiebe, E. C., Bitz, C. M., Duffy, P. B., Ewen, T. L., Fanning, A. F., Holland, M. M., MacFadyen, A., Matthews, H. D., Meissner, K. J., Saenko, O., Schmittner, A., Wang, H., and Yoshimori, M. (2001). The UVic earth system climate model: Model description, climatology, and applications to past, present and future climates. *Atmosphere-Ocean*, 39(4):361–428.
- Weber, T. and Deutsch, C. (2014). Local versus basin-scale limitation of marine nitrogen fixation. *Proceedings of the National Academy of Sciences of the United States of America*, 111(24):8741–8746.
- Wrightson, L. and Tagliabue, A. (2020). Quantifying the Impact of Climate Change on Marine Diazotrophy: Insights From Earth System Models. *Frontiers in Marine Science*, 7(July):1–9.
- Yao, W., Kvale, K. F., Achterberg, E., Koeve, W., and Oschlies, A. (2019). Hierarchy of calibrated global models reveals improved distributions and fluxes of biogeochemical tracers in models with explicit representation of iron. *Environmental Research Letters*, 14(11):114009.
- Zahariev, K., Christian, J. R., and Denman, K. L. (2008). Preindustrial, historical, and fertilization simulations using a global ocean carbon model with new parameterizations of iron limitation, calcification, and N<sub>2</sub> fixation. *Progress in Oceanography*, 77(1):56– 82.
- Zehr, J. P. and Capone, D. G. (2020). Changing perspectives in marine nitrogen fixation. *Science*, 368(6492):eaay9514.

A. Supporting Information for 'Hierarchy of calibrated global models reveals improved distributions and fluxes of biogeochemical tracers in models with explicit representation of iron'

## A.1. Calibration framework

Efforts into objective calibration of biogeochemical model parameters in prognostic ESMs are hindered by (a) the millennial time scales involved in whole ocean biogeochemical cycles, which require long model integration times and (b) the lack of suitable parameter calibration techniques. A solution to these obstacles exists in the form of offline methods that can use prescribed circulation and atmospheric boundary conditions, e.g. for seasonally cyclic conditions with the Transport Matrix Method (Khatiwala, 2007, TMM;) or data-assimilated, steady circulation (Primeau et al., 2013). The TMM is sufficiently fast to be combined with methods of parameter optimization. Recently, Kriest et al. (2017) used the TMM together with a Covariance Matrix Adaption Evolution Strategy (CMA-ES; Hansen, 2006) to calibrate a pre-selected subset of parameters of a global ocean biogeochemical model against available oceanic observations of  $PO_4$  and  $O_2$ .

An equilibrated circulation field is extracted from the UVic ESCM, creating an "offline" version of the UVic ESCM. The advantage of the offline model is a reduction of the model "spin-up" computing time by 2 orders of magnitude. This is made possible by two features of the TMM: the avoidance of time-consuming circulation computation and the ability to utilize parallel computing resources compared to the serial code of UVic ESCM 2.9. A comparison of the online and offline model performance is described in detail in Kvale et al. (2017). In the extraction of the circulation field, instead of FCT, the Up-Wind3 convection schema (Holland et al., 1998; Griffies et al., 2004) is used and the forcing is set to be pre-industrial setting (year 1800).

#### A.1.1. Calibration parameter selection

We calibrate three different iron modules using the same physical transport. The misfit function of  $J_T$  is applied for FeDyn, FeMask and NoFe. We also compare to the hand-tuned FeDyn from Nickelsen et al. (2015, denoted as FeDyn0). The number of biogeochemical parameters increases with increasing model complexity (from 30 in NoFe, to 34 in FeMask, and to 46 in FeDyn). We first conducted a parameter sensitivity study and chose six parameters for calibration ( $K_{Fe}^{P}max$ ,  $K_{Fe}^{D}$ ,  $g_{Z}^{0}$ ,  $\omega_{Det}^{i}$ ,  $K_{Fe}^{org}$  and  $R_{O:N}$  in Fe-Dyn; for definitions see Tab. 3.1, for formulations see Nickelsen et al., 2015; Keller et al., 2012). This parameter set contains some of the most uncertain (few or no measurements available) and sensitive biogeochemical model parameters. At the same time, it provides a large coverage of control in the different domains, e.g., iron, nitrogen, oxygen, phytoplankton, and zooplankton, of the model's biogeochemistry. For the other calibrations, we try to keep the parameter choices as similar to FeDyn as possible. Hence, for FeMask, we chose  $K_{Fe}^{\bar{D}}$ ,  $g_Z^0$ ,  $\omega_{Det}^i$  and  $R_{O:N}$  which are also presented in FeDyn. However, since FeMask does not represent the iron scavenging process, we left out  $K_{Fe}^{org}$  and replaced it with  $\mu_P^*$  (defined in Tab. 3.1).  $\mu_P^*$  determines the rate of microbial loop recycling, which influences the surface nutrient concentration, and thus impacts the nutrient limitation on phytoplankton growth. Also, the ordinary phytoplankton has a fixed iron uptake saturation  $K_{Fe}^{P}$ . In the FeDyn, Nickelsen et al. (2015) implemented the flexible iron saturation depending on biomass, which is described in three parameters, maximum iron saturation  $(K_{Fe}^{P}max)$ , minimum iron saturation  $(K_{Fe}^{P}min)$  and the biomass when iron saturation starts to increase  $(P_{max})$  (see equation 4-6, Nickelsen et al., 2015). For NoFe, two iron halfsaturation parameters are not in the model, which meant there was only macro-nutrient limitation control on phytoplankton growth. Hence, we added a, phytoplankton maximum growth rate at  $0^{\circ}C(a)$ , to the parameter set, so that the growth of phytoplankton is regulated in the calibration. The exact parameter set for each calibration and the parameter boundaries are listed in Tab. 3.1. The other parameters, which are not in the set, are unchanged from Nickelsen et al. (2015) and their values of are listed in Tab. A.1.

#### A.2. Calibration spin-up configuration

In our calibration, we run each calibration for 120 generations, or until the parameters converged. The number of individuals in each generation is 10, following the set up of Kriest et al. (2017). Each individual is integrated for 3000 model years with pre-industrial forcing (year 1800), and the vertical diffusion coefficient is increased from 0.15 to 0.43  $cm^2 s^{-1}$ , in order to mediate the strength reduction of the Meridional Overturning circulation (MOC) resulting from the advection scheme switch.

All calibrations were run at the Berlin facility of the North-German Supercomputing Alliance, on Intel Xeon IvyBridge computer nodes, 24 cores per node. For each individual run, 4 nodes were used for the parallel computation and it takes around 7 hours for a single 3000 year spin-up.

Calibrations converged within different numbers of generations due to the differences in model complexity. FeDyn converged within 120 generations and NoFe converged after only 100 generations. FeMask still had a small drift for  $g_Z^0$  approaching 120 generations. But the calibration was terminated after 120 generations, since the variance of  $J_T$  was at the magnitude of  $10^{-5}$  in the last 10 generations, which indicated the calibration results

from those generations are equally good. All calibration trajectories for the parameters are shown in the figure A.2,A.3 and A.4. We then selected the best parameter set (listed in Tab. 3.1), as the individual with the smallest  $J_T$  for each calibration. We next integrated the models using the optimal parameter sets with pre-industrial forcing for 9000 model years to achieve equilibrium, and diagnosed the different components of misfit, elemental fluxes, and oxygen minimum zone (OMZ) volumes for comparison.



**Figure A.1.:** Comparison between the annual basin-wide averaged nutrient and oxygen profiles in different basins from different calibrated models. Differences between the calibrated models are generally smaller than the differences between the models and observations. For example, in the Atlantic basin, the differences between calibrated models are much smaller (around 1  $mmol m^{-3}$  for NO<sub>3</sub>, 0.1  $mmol m^{-3}$  for PO<sub>4</sub> and 5  $mmol m^{-3}$  for O<sub>2</sub>) than the differences between models and observations (around 6.5  $mmol m^{-3}$  for NO<sub>3</sub>, 0.6  $mmol m^{-3}$  for PO<sub>4</sub> and 20  $mmol m^{-3}$  for O<sub>2</sub>). The similarities between the profiles of the three calibrated models suggest that they have approached some optimal limit for the given calibration setup. Additional improvements may require a better representation of the circulation and mixing processes.


**Figure A.2.:** Parameter trajectories in calibration of NoFe. The dashed straight lines denote the parameter calibration boundaries and the solid straight line denotes the uncalibrated parameter value. The bold black line is the average parameter value in each generation and the thin red lines shows the maximum and minimum value parameter value in each generation.



Figure A.3.: Parameter trajectories in calibration of FeMask. The notation follows figure A.3.



Figure A.4.: Parameter trajectories in calibration of FeDyn. The notation follows figure A.3.

Parameter	Description	Value	Units	NoFe	FeMask	FeDvn
1.Phytoplankton $(P_0, P_0)$	1					
	Maximum growth rate parameter at $0^{\circ}C$	0.6	$d^{-1}$	./		./
mp	$P_{\rm o}$ mortality rate	0.03	$d^{-1}$	v ./	v	v
те <sub>о</sub> и*	Microbial loop recycling rate	0.015	$d^{-1}$	v	v	v
μp	Diazotrophs' handican	0.015	u	v	v	v,
CD light limitation	Diazotrophis handicap	0.4		V	V	V
k	Light attenuation due to water	0.04	-1-1	/	/	
K <sub>W</sub> I₁	Light attenuation through phytoplankton	0.047	$m^{-1}$ (mm of N m <sup>-1</sup> ) <sup>-1</sup>	v	v	v
K <sub>C</sub>	Light attenuation through phytoplankton	5.0	m (mmor N m ) 1	v	V	v
	Eight attenuation unough sea ice & show	0.42	m	V,	v	V,
PAR	Fraction of photosynthetically active radiation	0.43	$\alpha u (\alpha) = 1$	$\checkmark$	V	V,
$\Theta_{min}$	Maximum Chil: Cratio, extreme fron limitation	0.01	$gChl(gC)^{-1}$		V	V,
$\theta_{max}$	Maximum Chi:C ratio, abundant iron	0.04	$gCnl(gC)^{-1}$	,	V,	$\checkmark$
α	Initial slope of the P-I curve	0.1	$(Wm^2)^{-1}d^{-1}$	$\checkmark$		,
$\alpha_{min}$	Minimum chlorophyll specific initial slope in PI-curve	18.4	$\mu g C (g Chl)^{-1} (W m^{-2})^{-1} s^{-1}$			V,
$\alpha_{max}$	Maximum chlorophyll specific initial slope in PI-curve	73.6	$\mu g C (g Chl)^{-1} (W m^{-2})^{-1} s^{-1}$			$\checkmark$
Macro-nutrient limitation			2		4	
$K_N$	Half saturation constant for N uptake	0.7	$mmol N m^{-3}$			
$R_{P:N}$	Molar P:N ratio	0.0625		$\checkmark$		$\checkmark$
$K_P$	Half saturation constant for P uptake	$K_N R_{P:N}$	$\mu mol m^{-3}$			
Iron limitation						
$K_{Fe}^P$	Half-saturation constant for $P_O$ iron limitation	0.1	$\mu mol Fe m^{-3}$			
K <sup>P</sup> <sub>Fe</sub> min	Minimum $K_{Fe}^P$	0.04	$\mu mol Fe m^{-3}$			$\checkmark$
$K_{Fe}^{P}max$	Maximum $K_{Fe}^{P}$	0.4	$\mu mol Fe m^{-3}$			$\checkmark$
P <sub>max</sub>	$P_O$ biomass above which $K_{Fe}^P$ increases	0.15	$mmol N m^{-3}$			$\checkmark$
$K^D_{Fe}$	Half-saturation constant for Diaz Iron limitation	0.1	$\mu mol Fe m^{-3}$			
R <sub>Fe:N</sub>	Molar Fe:N ratio for iron uptake	66.25	$\mu mol Fe (mol N)^{-1}$			
2. Zooplankton (Z)	*					·
	Maximum grazing rate at $0^{\circ}C$	0.4	$d^{-1}$	1/	1	
82 9 <sup>e</sup> 7	Z growth efficiency	0.4	u	v	v v	v
82 V	Z assimilation efficiency	0.7		v ./	v	v ./
1 117	Z quadratic mortality	0.06	$(mmol N m^{-3})^{-2} d^{-1}$	v ./	V.	v ./
K <sub>C</sub>	half-saturation constant for Z grazing	0.00	$mmol N m^{-3}$	v	v	v ./
Wa	Z preference for $P_{\rm o}$	0.15	minoritim	v	v	v
$\Psi P_O$	Z preference for $P_{\rm D}$	0.5		v	v	v,
$\Psi P_D$	Z preference for other Z	0.1		v	v	v
ΨZ Wp.	Z preference for <i>Det</i>	0.3		v	v	v
$\frac{\Psi Det}{2}$		0.5		V	V	<u>v</u>
3. Detritus (Det)	D to a line in the OSC	0.055		/		
$\mu_{Det}$	Det remineralization rate at 0 C	0.055	1-1	V,	V	V,
$\omega_{Det}$	Sinking speed of <i>Del</i> at surface	14		V,	V	V,
$\omega_{Det}$	Depth dependent sinking speed slope	0.06	a ·	$\checkmark$	$\checkmark$	$\checkmark$
Carbonate	Malan C. Martin	6 625		/	/	
R <sub>C:N</sub>	CoCO2 over popularil BOC production ratio	0.025		v	v	V
R <sub>CaCO3</sub> :POC	CaCO3 over nonaigar POC production ratio	6500		v	V	V
D <sub>CaCO3</sub>	CaCO3 remineralisation e-folding depth	6500	m	V		<u></u>
4.Iron (Fe)			2			
$L_T$	Total ligand concentration	1	$\mu$ mol lig m <sup>-3</sup>			
K <sub>FeL</sub>	Fe-ligand stability constant	105.5	$(mmol  lig  m^{-3})^{-1}$			
$R_{Fe:P_{sed}}$	Molar <i>Fe</i> : <i>P</i> for sedimentary iron source	0.004				
$O_{2min}$	Minimum O2 concentration for aerobic oxidation	5	$mmol O_2 m^{-3}$	$\checkmark$	$\checkmark$	
$K_{Fe}^{org}$	Organic-matter dependent scavenging rate	0.45				$\checkmark$
$K_{Fe}^{O'8}s$	Organic-matter particle scaling for scavenging	0.58				
K <sub>Fe</sub>	Fe colloidal production and precipitation rate	0.005	<i>d</i> <sup>-1</sup>			
5. Oxygen ( <i>O</i> <sub>2</sub> )						
R <sub>O:N</sub>	Molar O: N ratio	10		$\checkmark$		$\overline{}$
6. Temperature dependency						<u> </u>
<i>b</i>	Scaling factor b for temperature dependent process	1.066		1		
с	Scaling factor c for temperature dependent process	1.0	$(^{\circ}C)^{-1}$	v v	Ň	Ň

**Table A.1.:** The parameters chosen for each calibration, their observational value and calibration boundary settings.

## B. Supporting Information for 'Simulated future trends in marine nitrogen fixation are sensitive to model iron implementation'

## B.1. Nutrient and light limitation on phytoplankton growth

The nutrient limitation factor is defined in the model as a function of nutrient concentration and phytoplankton nutrient uptake half-saturation rate:

$$X_{Lim(OorD)} = \frac{C_X}{C_X + k_{X(OorD)}}$$
(B.1)

, X is NO<sub>3</sub>, PO<sub>4</sub> or *Fe*,  $C_X$  is the concentration of X and  $k_{X(OorD)}$ , is the phytoplankton half-saturation of the uptake rate for X. Light limitation is derived from Keller et al. (2012, [for NoFe and FeMask) and Nickelsen et al. (2015, for FeDyn) as:

$$L_{lim(OorD)} = \frac{\alpha_{(OorD)}I}{\sqrt{(J_{max(OorD)}^2 + (\alpha_{(OorD)}I)^2)}}$$
(B.2)

, where  $\alpha_{(OorD)}$  is the initial slope of the photosynthesis-irradiance curve (in FeDyn,  $\alpha_{(OorD)}$  is dependent on  $Fe_{Lim(OorD)}$ , details see Nickelsen et al., 2015), *I* is the in situ irradiance and  $J_{max(OorD)}$  is the in situ maximum potential growth rate.

$$J_{max(O or D)} = a \cdot F_{temp(O or D)} \tag{B.3}$$

*a* is the phytoplankton growth rate constant at 0 °C.  $F_{temp(OorD)}$  is a temperature factor for biological process.

$$F_{tempO} = exp(T/T_b) \tag{B.4}$$

$$F_{tempD} = Cd \cdot max(0, a(exp(T/Tb) - 2.61))$$
(B.5)

*T* is the situ temperature and  $T_b$  is the e-folding temperature of biological rates,  $C_d$  is a handicap for diazotroph growth and the 2.61 temperature scaling is to inhibit diazotrophs growth below 15 °C.

The iron is not incorporated in the nutrient minimum function but as a factor that directly governs the growth rate. The in situ growth rate of ordinary phytoplankton  $(J_O)$  and diazotrophs  $(J_D)$  are defined as:

$$J_O = min(NO3_{limO}, PO4_{limO}, L_{limO}) \cdot Fe_{LimO} \cdot a \cdot F_{tempO}$$
(B.6)

97



**Figure B.1.:** The most limiting growth factor for ordinary phytoplankton in all models in the years 1800 (upper) and 2100 (lower).

$$J_D = min(PO4_{limD}, L_{limD}) \cdot Fe_{LimD} \cdot a \cdot F_{tempD}$$
(B.7)

More details of model equations can be found in Keller et al. (2012, FeMask and NoFe) and Nickelsen et al. (2015, FeDyn).



**Figure B.2.:** Annually averaged total NPP, export, and oxygen concentration (depth 350 *m*) in the Eastern Tropic Pacific. Color shading represents the difference between year 2100 and 1800 and contours show simulated values in year 2100. The unit for NPP and export production is  $gCm^{-2}yr^{-1}$  and the unit for O<sub>2</sub> concentration is  $mmol m^{-3}$ . The interval between neighboring lines are 50  $gCm^{-2}yr^{-1}$  for NPP, 2  $gCm^{-2}yr^{-1}$  for export, and 5  $mmol m^{-3}$  for O<sub>2</sub>.

Parameter	Description	NoFe	FeMask	FeDyn	Units			
Turied Description Note Toblask Tebyn Ullis								
1. Prytopiankton $(r_0, r_D)$								
а	Maximum growth rate parameter at 0°C	0.698	0.0	0.0	a			
$m_{P_O}$	$P_O$ mortality rate	0.03	0.03	0.03				
$\mu_P^*$	Microbial loop constant	0.0157	0.0012	0.015	$d^{-1}$			
$c_D$ Diazotrophs handicap $0.4$ $0.4$ $0.4$								
light limitati	on				1			
$k_w$	Light attenuation due to water	0.04	0.04	0.04	$m^{-1}$			
$k_c$	Light attenuation through phytoplankton	0.047	0.047	0.047	$m^{-1} (mmol N m^{-1})^{-1}$			
$k_I$	Light attenuation through sea ice & snow	5.0	5.0	5.0	$m^{-1}$			
PAR	Fraction of photosynthetically active radiation	0.43	0.43	0.43				
$\theta_{min}$	Maximum Chl:C ratio, extreme iron limitation		0.01	0.01	$gChl(gC)^{-1}$			
$\theta_{max}$	Maximum Chl:C ratio, abundant iron		0.04	0.04	$gChl(gC)^{-1}$			
α	Initial slope of the P-I curve		0.1	0.1	$(W m^{-2})^{-1} d^{-1}$			
$\alpha_{min}$	Minimum chlorophyll specific initial slope in PI-curve			18.4	$\mu g C (g Chl)^{-1} (W m^{-2})^{-1} s^{-1}$			
$\alpha_{max}$	Maximum chlorophyll specific initial slope in PI-curve			73.6	$\mu g C (g Chl)^{-1} (W m^{-2})^{-1} s^{-1}$			
Macro-nutrie	ent limitation							
KN	Half saturation constant for N uptake	0.7	0.7	0.7	$mmol N m^{-3}$			
RDN	Molar P·N ratio	0.0625	0.0625	0.0625				
K <sub>P</sub>	Half saturation constant for P untake	Ky Rp y	Ky Rpy	Ky Rpy	$\mu mol m^{-3}$			
Iron limitati		K <sub>N</sub> KP:N	M <sub>N</sub> MP:N	$\mathbf{M}_N \mathbf{M}_{P:N}$	μποι π			
KP	Half saturation constant for $P$ iron limitation		0.05 /		umol Fam <sup>-3</sup>			
K <sub>Fe</sub>	Half-saturation constant for $F_0$ from minitation		0.05	0.04	$\mu mot Fem$			
K <sub>Fe</sub> min V <sup>P</sup>	Minimum $K_{Fe}$			0.04	$\mu$ mol Fem <sup>-3</sup>			
$K_{Fe}^{i}max$	Maximum $K_{Fe}$			0.406	µmol Fem <sup>3</sup>			
Pmax	$P_O$ biomass above which $K_{Fe}^r$ increases			0.15	$mmol N m^{-3}$			
$K_{Fe}^D$	Half-saturation constant for Diaz Iron limitation		0.406√	0.377√	$\mu mol Fem^{-3}$			
$R_{Fe:N}$	Molar Fe:N ratio for iron uptake			66.25	$\mu mol Fe (mol N)^{-1}$			
2. Zooplank	ton (Z)							
$g_Z^0$	Maximum grazing rate at 0°C	0.4	0.4	0.4	$d^{-1}$			
$g_Z^e$	Z growth efficiency	1.282	0.668	0.567				
γ	Z assimilation efficiency	0.7	0.7	0.7				
mZ	Z quadratic mortality	0.06	0.06	0.06	$(mmol N m^{-3})^{-2} d^{-1}$			
KGraze	half-saturation constant for Z grazing	0.15	0.15	0.15	$mmol N m^{-3}$			
WPo	Z preference for $P_{\Omega}$	0.3	0.3	0.3				
$W_{P_{D}}$	Z preference for $P_{\rm D}$	0.1	0.1	0.1				
W7	Z preference for other Z	0.3	0.3	0.3				
Ψ2 ₩D-4	Z preference for <i>Det</i>	0.3	0.3	0.3				
<u></u>	Det remineralization rate at 0°C	0.055	0.055	0.055				
$\mu_{Det}$	Similar arread of Dat at surface	0.035	0.033	0.035				
$\omega_{\tilde{D}et}$	Sinking speed of <i>Det</i> at surface	14	14	14	ma			
$\omega_{Det}$	Depth dependent sinking speed slope	0.065	0.062	0.060	a			
Carbonate		( ( ) =	( ( ) =					
$R_{C:N}$	Molar C : N ratio	6.625	6.625	6.625				
$R_{CaCO_3:POC}$	CaCO3 over nonalgal POC production ratio	0.03	0.03	0.03				
$D_{CaCO_3}$	CaCO3 remineralisation e-folding depth	6500	6500	6500	m			
4.Iron ( <i>Fe</i> )								
$L_T$	Total ligand concentration			1	$\mu$ mol lig m <sup>-3</sup>			
$K_{FeL}$	Fe-ligand stability constant			$10^{5.5}$	$(mmol  lig  m^{-3})^{-1}$			
R <sub>Fe</sub> :Paul	Molar Fe : P for sedimentary iron source			0.004				
$O_2 min$	Minimum O2 concentration for aerobic oxidation	5	5	5	$mmol O_2 m^{-3}$			
K <sup>org</sup>	Organic-matter dependent scavenging rate			0.427	-			
$K_{F_{2}}^{re}s$	Organic-matter particle scaling for scavenging			0.58				
K <sup>col</sup>	Fe colloidal production and precipitation rate			0.005	$d^{-1}$			
5 Overage (				5.005				
J. Oxygen (	Molor Q + N rotio	10.420 /	0.541 /	10.502 /				
K <sub>O:N</sub>		10.439	9.541	10.502				
6. Temperati	ure dependency							
b	Scaling factor b for temperature dependent process	1.066	1.066	1.066				
С	Scaling factor c for temperature dependent process	1.0	1.0	1.0	$(^{\circ}C)^{-1}$			

**Table B.1.:** The biogeochemical parameter values for each model. Summarized from Yao et al. (2019).

 $\sqrt{}$ : parameter value optimised for the model.

## C. Acknowledgements

First, I want to thank Andreas Oschlies for the privilege to work in this inspiring institute and to be under his supervision. He gave me the opportunity to visit conferences, workshops and summer schools, which expanded my horizon and connected me with so many interesting people. I want to thank him for his friendly demeanor, insightful advice, and his unwavering support throughout my Ph.D. period.

I also want to thank my second supervisor Eric Achterberg, who took a chance on me. I thank him for his contagious high-spirit even in the gloomiest days in Kiel, for his wonderful sense of humor, and for his direct and honest opinions.

I want to thank Wolfgang Koeve, who believed in me even when I did not believe in myself. I thank him for his patience in nurturing me and training me to become a better scientist. I always appreciate the discussions with him, which helps me to put things into perspective. I just simply enjoy the time I spent working with him.

I want to thank Karin F. Kvale, who has supported me greatly both on a professional and personal level. I thank her for her whimsical ideas, her flowery language, and her dedication.

I would like to thank Erin Bertrand, Katja Fennel, and Alessandro Tagliabue for taking part in my supervising committee. I thank them for the time they spent on my committee meetings, for their feedback on my work, and the discussions through out the years.

I also want to thank the members in the biogeochemical modelling group, who are always friendly and helpful.

A special thank you to the HOSST-TOSST program (Christel, Christian, David), who provided me with founding and made the last five years possible.

Other than that I would like to thank my fellow students, especially Fabian, Kirsten, Eva, Tatum, Meghan, Florian, Jiajun, and Tianfei, with whom I shared some of the most precious memories for the last few years and with whom I wish to share more in the future.

A special thanks to Allison Chua, who is not only an inspiring role model for me but also

a sister i wish I had.

Finally, I want to thank my friends and family for the moral support. This thesis is dedicated to Wan Li, a courageous, free-spirited woman and my patient, loving mother.



## D. Declaration

I hereby declare that, apart from the supervisor's guidance, the content and design of the thesis "Implications of Iron Model Complexity for the Projection of Global Biogeochemical Cycles - From the Preindustrial Era and into the Future" is all my own work. Apart from the referenced sources and advice from my supervisors, there is no other source of information for this thesis. This thesis has been prepared subject to the Rules of Good Scientific Practice of the German Research Foundation. The thesis has not been submitted either partially or wholly as a part of a doctoral degree to another examining body and is my first and only doctoral procedure. I have not been deprived of an academic degree. I understand that this thesis will be passed on to the library of GEOMAR Helmholtz Centre for Ocean Research Kiel and the library of Christian-Albrechts-University in Kiel.

Kiel, April 2021

(Wanxuan Yao)