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

List of Tables

List of Figures

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 pre-industrial 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_2-Konzentrationspfades, der mit einem „Business-as-usual"-CO_2-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_2. 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 nitrogen-depleted 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).

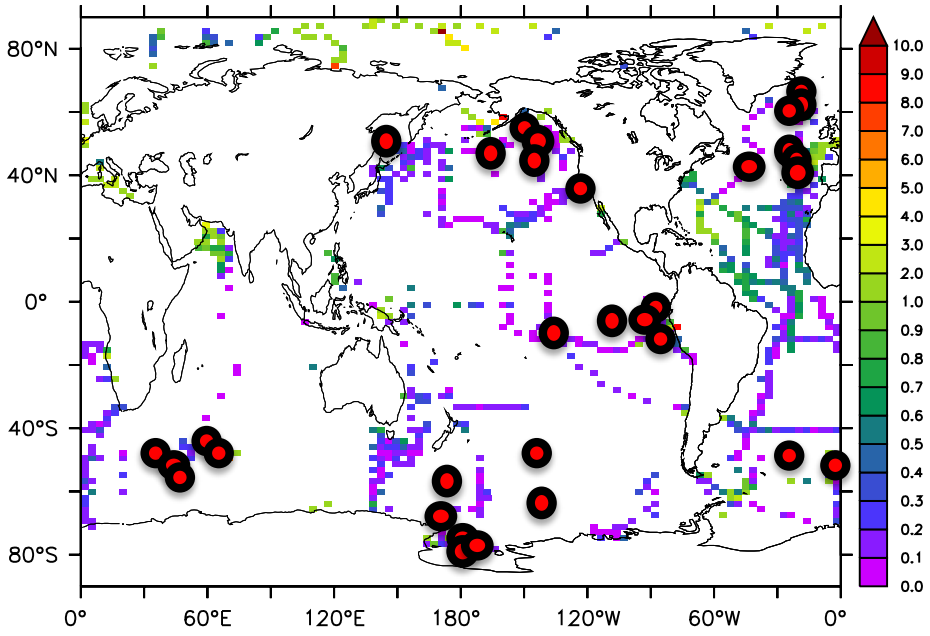

Figure 1.1.: *Observed dissolved iron concentrations in the surface ocean (tiles; $\mu mol\, m^{-3}$) and locations that are primarily iron limited (red points within black circles). The iron concentrations are compiled from both the 2017 GEOTRACES intermediate data product (Schlitzer et al., 2018) and (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 $mmol\, mol^{-1}$ (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

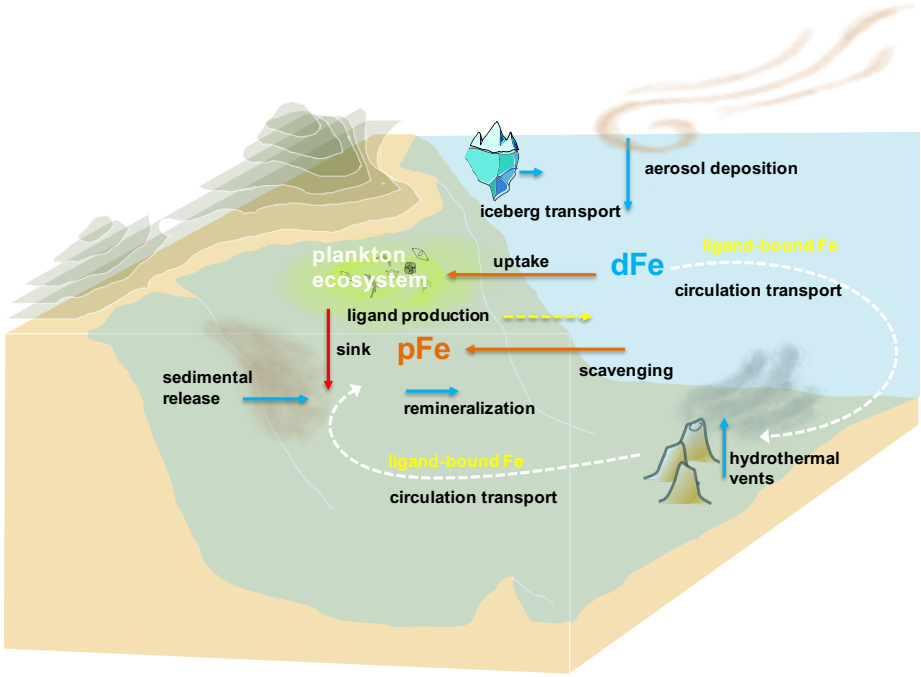

Figure 1.3.: *The iron cycle in the ocean. This is a representation of the marine iron cycle adapted from the review on marine iron cycling by Tagliabue et al. (2017). The sources of iron in the ocean are aerosol deposition, sedimentary release, iceberg transport, and hydrothermal flux. Phytoplankton 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 GEOTRACES 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 advective-diffusive 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 $^\circ$ latitude $\times$ 3.6 $^\circ$ 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-

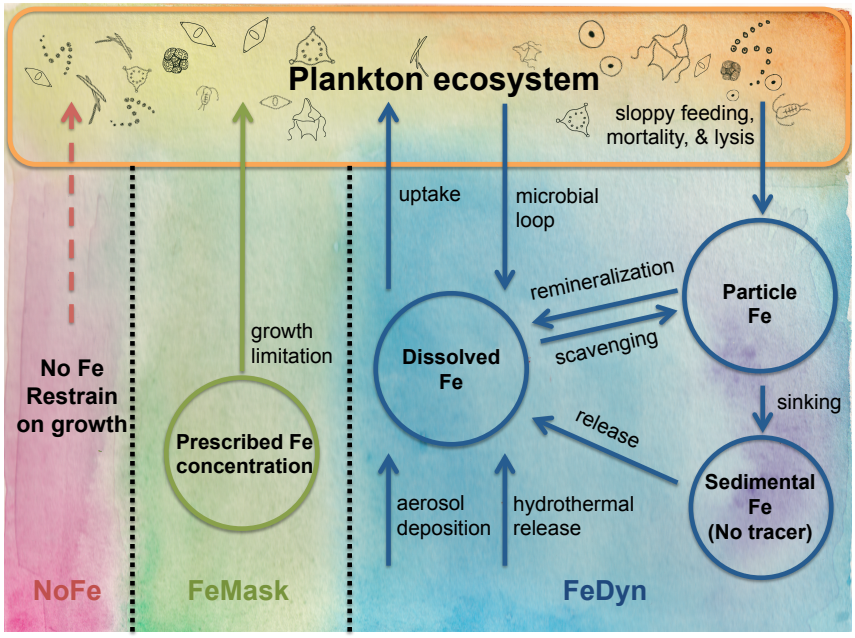

Figure 1.4.: *Three different implementations of iron modules in a biogeochemistry model. The complexity increases from left to right (NoFe < FeMask < FeDyn).*

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 stoichiometric 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.