A Data Assimilating Model of the Microbial Ecosystem in the Open Ocean

By Yawei Luo

B.A., Ocean University of China, 2000

M.S., Peking University, 2003

A Dissertation Submitted in Partial Fulfillment of the

Requirements for the Degree of Doctor of Philosophy

in the Program in BioMed: Ecology and Evolutionary Biology at Brown University

Providence, Rhode Island

May 2010

This dissertation by Yawei Luo is accepted in its present form

by the Department of Ecology and Evolutionary Biology as satisfying the

dissertation requirement for the degree of Doctor of Philosophy

Date_____ _____

Hugh W. Ducklow, Advisor

Recommended to the Graduate Council

Date_____ _____

Marjorie A.M. Friedrichs, Reader

Date_____ _____

Warren L. Prell, Reader

Date_____ _____

Joseph J. Vallino, Reader

Approved By the Graduate Council

Date_____ _____

Sheila Bonde, Dean of the Graduate School

Curriculum Vitae

Yawei Luo was born May 17, 1978 in the city of Leshan, Sichuan Province, China. He grew up in the city until he attended Ocean University of China in Qingdao, Shandong Province in September 1996, and started his journey in marine science. In June 2000, he completed his Bachelor's degree from studying fishery resources and was also trained with computer science as the second major. In the following September, he moved to Peking Univeristy in Beijing, China, and attended a Master program in environmental science and worked with Drs. Huiding Wu, Baozhang Liu and Xuhui Cai. He did a wind-tunnel experiment on automobile pollution in Macau, developed an application software for monitoring and forecasting sea ice from the images of the first ocean satellite of China, and conducted a marine ecosystem modeling in Bohai Sea. After he completed his Master's degree in June 2003, he relocated in the following August to Gloucester Point, VA, USA to attend a Ph.D. program in the School of Marine Science in the College of William and Mary and worked with Dr. Hugh Ducklow. In 2007, he transferred to Brown University – Marine Biological Laboratory (MBL) Graduate Program in Woods Hole, MA, USA to continue his program with Dr. Ducklow who was taking a new job at Marine Biological Laboratory (MBL). His work in this program, a data assimilating model of marine ecosystem, was presented in this dissertation. In August 2009, he successfully defended his Ph.D.

PUBLICATIONS

Luo, Y., Ducklow, H. W., Friedrichs, M. A. M., Doney, S. C., Church, M. J., Prell, W., Vallino, J. J (*submitted*) Heterotrophic microbial dynamics in the open ocean revealed by data assimilation modeling. Aquatic Microbial Ecology.

Luo, Y., Ducklow, H. W., Friedrichs, M. A. M., Doney, S. C. (*in prep.*) A 15-Year Data Assimilating Model of the Marine Microbial Ecosystem at Station ALOHA, Hawaii.

Luo, Y., Ducklow, H. W. (*in prep*) Constraining a marine ecosystem model with heterotrophic bacterial production measurements

Gasol, J.M., Pinhassi, J., Alonso-Sáez, L., Ducklow, H., Herndl, G.J., Koblízek, M., Labrenz, M., **Luo, Y.**, Morán, X.A.G., Reinthaler, T., Simon, M. (2008) Towards a better understanding of microbial carbon flux in the sea. Aquatic Microbial Ecology 53:21-38.

Luo, Y., Wu, H., Zhang, Y., Sun, C., Liu, Y. (2004) Application of the HY-1 satellite to sea ice monitoring and forecasting. Acta Oceanologica Sinica 23:251-266.

ACKNOWLEDGEMENTS

The fact this dissertation has come to fruition has only been possible due to the support, help and encouragement of numerous individuals. First and foremost, I wish to thank my doctoral advisor, the world-leading marine microbiologist, Dr. Hugh Ducklow. It is an unbeatable opportunity to work so closely with him, as he is such a supportive, creative and motivated advisor one can ever imagine. He is always ready to help, while respects my opinions and encourages independent research. His personality and enthusiasm on science has created a model for my career and will benefit my life greatly in the future.

I would like to thank Dr. Marjorie Friedrichs. This dissertation can never be done without her advices and mentorship. Marjy is more than a committee member. She introduces me the data assimilation, which is a fundamental part of this dissertation, and keeps greatly supporting my work and solving problems in the past three years. Her advices are always very creative and helpful. I feel so lucky to work with her.

I would also like to give my special thanks to Dr. Scott Doney. He has been discussing this dissertation with me in the past years. Many creative ideas came from the discussions. He also kindly reviewed this manuscript and provided many important comments that have greatly improved the quality of this dissertation.

I am grateful to my committee members Drs. Warren Prell and Joseph Vallino (Brown-MBL period), and Drs. Courtney Harris, Eileen Hofmann, Deborah Steinberg and Kam Tang (VIMS period) for their critical reviews and interests in my work.

There have been many students and scientists helped me a lot for both my work and life. Jeff Dusenberry constructed the original codes for the data assimilation framework.

Matthew Church kindly provided bacterial production data and had great discussions with my work. VIMS, MBL and Brown University have great environment to support research and life for students, especially for me as an international student. Here are some people helped me greatly: Iris Anderson, Gina Burrell, Fonda Powell, Sue Presson, Sarah Hamrick, Steve Sechrist, Jerry Melillo, John Hobbie, Ed Rastetter, Jim Tang, Kelly Holzworth, Debbie Scanlon, Mary Ann Seifert, Joshua Hamilton, David Rand, Carol Casper, Jamie Kendrioski, Elke Breker, and many others I cannot list them all. Brown-MBL students Priya Dutta, Gillian Galford, Yuko Hasegawa, Kristen Myers, Erica Lasek-Nesselquist, Alex Valm and I always support each other through the Journal Club. Of course my lab members have supported me most and make me feel the lab is a real family: Helen Quinby, Matthew Erickson, Kristen Myers, Heidi Geisz, Xelu Morán, Mirko Lunau and Maggie Waldron.

The computing resource is provided by the Sciclone cluster at the College of William and Mary. Tom Crockett works hard to maintain the system and is always very helpful to solve problems.

I wish to thank all the crews and scientists who collected and analyzed data during U.S. JGOFS cruises in Arabian Sea, Equatorial Pacific and Hawaii Ocean Time-series. I understand there was very hard work behind each single datum.

Special thanks to all the crews and scientists at LTER Jan 2006 cruise, the extraordinary experience I will never forget.

My family is always being with me through my entire career. I can never thank enough for everything you have done for me. This dissertation can never be done without your support and love.

Table of Contents

Chapter 1. Introduction 1

Overview 2

Ecological Backgroud: Consideration for Modeling 3

Heterotrophic Bacteria, DOM and Nutrient Regeneration 4

Nitrogen Fixation 10

Flexible C:N:P Stoichiometry 14

Data Assimilation 16

Study Site 18

Scientific Questions 20

Dissertation Structure 21

References 24

Chapter 2. A Marine Ecosystem Model with Microbial Dynamics and Its Application to a Data Assimilative Testbed for Open Ocean 39

Abstract 40

1. Introduction 42

2. Ecosystem Model Description 44

2.1. State Variables 44

2.2. Parameters 45

2.3. Temperature Effect 45

2.4. Phytoplankton (PHY) Processes 46

2.5. Diazotrophic Phytoplankton Processes 52

2.6. Bacterial Processes 55

2.7. Zooplankton (PRT and MZ) Processes 62

2.8. Other Processes (DET, dissolved inorganic nutrients & DOM) 63

3. Data Assimilative Testbed 64

3.1. Forward Model Setup 66

3.2. Assimilative Data and Cost Function 74

3.3. Adjoint Model 79

3.4. Optimization Procedure: Limited-memory Quasi-Newton Method 79

3.5. Selection of Optimizable Parameters and Uncertainty Analysis 80

3.6. Default Parameter Values and Optimization 82

Appendix A: Ecosystem Model Equations 84

Appendix B. Adjoint Model Construction 100

Appendix C. Limited-memory Quasi-Newton Method 104

References 107

Chapter 3. Heterotrophic Microbial Dynamics in the Open Ocean Revealed by Data Assimilative Modeling 126

Abstract 127

1. Introduction 129

2. Methods 132

2.1. The Ecosystem Model 132

2.2. Study Sites and Modeling Framework 135

2.3. Data Assimilation 137

3. Results 142

3.1. Fit of Modeling Results to Observations 142

3.2. Bacterial Production 144

3.3. DOM Production and Turnover Time 144

3.4. Nitrogen Budget 145

4. Discussion 147

4.1. Model Schemes of Microbial Dynamics 147

4.2. Dynamics of Bacterial Activity 150

4.3. Fate of Bacterial Production 153

4.4. Export in form of refractory DOM 155

5. Conclusions 157

References 160

Chapter 4. A 15-Year Data Assimilating Model of the Marine Microbial Ecosystem at Station ALOHA, Hawaii. 198

Abstract 199

1. INTRODUCTOIN 200

2. METHODS 202

3. Results 204

3.1. Model Optimization 204

3.2. Fit of Specific State Variables and Rates 206

3.3. Average Modeled Results in the 15-year Period 208

3.4. Interannual Variation of Primary Production 210

4. Discussion 220

4.1. Model explanations for the interannual primary production increase 220

4.2. Increase physical DIN input 222

4.3. N_2 fixation and the related Low N Recycling Efficiency 224

5. Conclusion 226

References 227

Chapter 5. Constraining a Marine Ecosystem Model with Heterotrophic Bacterial Production Measurements 271

Abstract 272

1. Introduction 273

2. Methods 275

3. Results and Discussion 276

3.1. Experiment 1: No BP assimilated 276

3.2. Experiment 2: High and Low BP assimilated 279

3.3. Necessity of BP measurements? 283

3.4. Uncertainty of BP 285

4. Conclusion 287

References 289

Chapter 6. Summaries and Conclusions 303

1. Biological Model 304

2. Data Assimilating Model 305

3. Findings 307

References 310

List of Tables

Chapter 2

Table 1. List of the state variables 119

Table 2. Default model parameter values 120

Table 3. Assimilative data availability 125

Chapter 3

Table 1. Average observations and modeled results 170

Table 2. Heterotrophic bacterial related results 173

Table 3. Fate of primary carbon production 174

Table 4. Turnover time of bacteria and DOM 175

Table 5. Fate of bacterial production 176

Table 6. Cost function reduction 177

Table 7. Default and optimized model parameters 178

Chapter 4

Table 1. Default and optimized model parameters 232

Table 2. Observed primary production in different periods 235

Chapter 5

Table 1. List of the parameter names of Figure 1 292

Table 2. Cost function values with or without BP 293

Table 3. BP to PP ratio and changed conversion factors 294

Table 4. Key modeled results with different BP assimilted 295

List of Figures

Chapter 1

Figure 1. Map of Study Sites 38

Chapter 2

Figure 1. Model structure 118

Chapter 3

Figure 1. Average modeled stocks and fluxes 190

Figure 2. Sources for DOM production 193

Figure 3. Nitrogen recycle and input 194

Figure 4. Labile and semilabile DOC production and consumption 195

Figure 5. Bacterial growth efficiency vs. bacterial production 196

Figure 6. Primary production and bacterial production 197

Chapter 4

Figure 1. Reduction of cost function 241

Figure 2. Comparison of depth-integrated observations and model results 242

Figure 3. Modeled average results 251

Figure 4. Observed and modeled interannual increase of primary production 253

Figure 5. Observed and modeled primary production 254

Figure 6. Annual primary production increase for three phytoplankton groups 255

Figure 7. Primary production increase rate at each vertical layer 256

Figure 8. Primary production increase rates at each month 257

Figure 9. Carbon and nitrogen components of primary production increase in 0 – 20 m 258

Figure 10. Annual photosynthetically active radiation..... 259

Figure 11. Annual nitrate and ammonium in 0 – 20 m..... 260

Figure 12. Change of nitrate and ammonium fluxes in 0 – 20m..... 261

Figure 13. Physical DIN transport and nitrogen fixation in 0 – 20 m..... 262

Figure 14. Nitrogen recycling efficiency through different phytoplankton groups . 263

Figure 15. Change of protozoan and metazoan fluxes in 0 – 20 m..... 264

Figure 16. Change of bacterial and DON fluxes in 0 – 20 m..... 265

Figure 17. Variability of semilabile DON production and consumption in 0 – 20 m 266

Figure 18. Carbon and nitrogen components of primary production increase in 50 – 70 m..... 267

Figure 19. Change of nitrate and ammonium fluxes in 50 – 70 m..... 268

Figure 20. Change of bacterial and DON fluxes in 50 – 70 m..... 269

Figure 21. Mixed layer depth 270

Chapter 5

Figure 1. Optimized parameters and confidence intervals with and without BP assimilated..... 298

Figure 2. Average modeled results at EQP without BP assimilated 299

Figure 3. Optimized parameters and confidence intervals with low, default and high BP assimilated..... 301

Figure 4. Change of cost function with low and high BP assimilated..... 302

Chapter 1. Introduction

Overview

This Ph.D. dissertation conducts 1-dimensional (1-D) ecosystem modeling of the oligotrophic open ocean. The objectives of the dissertation are (1) to study the dynamics of heterotrophic bacteria and dissolved organic matter (DOM) and their influences on the ocean carbon cycle and nutrient regeneration; (2) to study the dynamics of dinitrogen (N_2) fixation by autotrophic cyanobacteria and its role in enhancing primary production and influencing the ecosystem structure; (3) to reveal the temporal variability of heterotrophic bacteria/DOM and N_2 fixation dynamics as well as other major components of the ecosystem; (4) to illuminate why a flexible stoichiometry of carbon, nitrogen and phosphorus is necessary to solve the above issues and improve model performance; and (5) to reveal the importance of data assimilation in marine ecosystem modeling and demonstrate how to use observations to constrain models.

The major study site for this dissertation is Station ALOHA of the Hawaii Ocean Time-series (HOT) program. This site is located in the North Pacific Subtropical Gyre (NPSG), which is the largest ocean circulation pattern on the earth (Sverdrup et al. 1946). The station is selected because the condition of the site is expected to represent the vast area of NPSG and the circulation pattern indicates low horizontal transport, which can facilitate 1-D modeling. Also its oligotrophic condition is ideal for a study of heterotrophic bacteria and N_2 fixation dynamics.

An ecosystem model is developed to address the objectives. The long-term observations at Station ALOHA provide numerous data, which are crucial for ecosystem modeling. To take advantage of the observations, a 1-D data assimilative testbed is set up for the station to accommodate the ecosystem model and to minimize the misfits between

the model and observations by objectively optimizing the model parameters.

Ecological Backgroud: Consideration for Modeling

The gas exchange between air and sea is a purely physical-chemical process. The partial pressure of CO_2 (pCO_2) in the upper ocean is one of the key factors determining the magnitude and direction of air-sea CO_2 flux. The net CO_2 flux is from air to the ocean when pCO_2 in the air is higher than the upper ocean, and vice versa. The larger the difference of pCO_2 between the air and the ocean (ΔpCO_2), the higher the transfer rate. The biological activities of marine ecosystems can significantly change pCO_2 in the ocean and therefore air-sea CO_2 flux in the upper ocean (Carlson 2001, Ducklow et al. 2001). Net primary production converts CO_2 into organic matter, some of which is exported to depth, where the organic matter is respired back to CO_2. This process, called the biological pump, reduces the dissolved inorganic carbon (DIC) level in the upper ocean and enhances the CO_2 flux into the ocean. One way to evaluate this process is to study the magnitude of exported carbon including both particulate organic carbon (POC) and dissolved organic carbon (DOC). We have estimates of POC export from sediment traps (Berelson 2001, Buesseler et al. 2007, Honjo et al. 2008), DOC export from its concentration gradient (Carlson et al. 1994, Ducklow et al. 1995, Hansell 2002) or total organic carbon export indirectly from mass balances of dissolved oxygen, inorganic carbon and organic carbon (Emerson et al. 1997). The mechanisms controlling carbon export are not well understood. In this dissertation the major emphasis is on the roles of microbes in the marine carbon cycle. In the following section I will discuss why the

dynamics of heterotrophic bacteria and N_2 fixation are important to address the ocean carbon cycle.

Heterotrophic Bacteria, DOM and Nutrient Regeneration

About 50% of primary production by phytoplankton in oceanic ecosystems (Ducklow & Carlson 1992, Williams 2000) is released as dissolved organic matter (DOM) through complex biotic and abiotic processes (Carlson 2002, Kirchman 2004). DOM is then mostly utilized by heterotrophic *Bacteria* and possibly also *Archaea*. These two groups (termed bacteria hereafter) can be grazed by protists and then by higher trophic levels; therefore this otherwise lost production can be partially recycled back to the food web. This process is termed the "microbial loop" (Azam et al. 1983). Most of the organic carbon production metabolized by bacteria is respired to CO_2 (mean value 73% ~ 91%; Anderson & Ducklow 2001), and the average ratio of bacterial production (BP) to primary production (PP) in the open oceans is ~15% (Ducklow 2000).

DOM is one of the largest exchangeable carbon reservoirs on earth (Hedges 1992). The dissolved organic carbon (DOC) pool is estimated to be 0.60×10^{18} g Carbon (Williams & Druffel 1987), a value comparable to the mass of C in atmospheric CO_2 in size (ca. 0.66×10^{18} ton Carbon) (Olson et al. 1985). Thus as the major consumer of the DOM pool, bacteria are a key component in the global carbon cycle. A small change of bacterial activities may have significant influence on air-sea carbon exchange.

Bacteria are also a key component of marine nutrient recycling in the euphotic zone. Dissolved organic nitrogen and phosphorus (DON and DOP) must be processed mainly by bacteria to be remineralized to inorganic forms (Kirchman 2000, Landry 2001). The

ratio of PP supported by external nutrients to total PP (supported by external plus recycled nutrients), commonly termed the "f-ratio" (Eppley & Peterson 1979), is generally low in the open ocean (e.g., 11-36% in Anderson & Ducklow 2001). Thus PP in the open ocean depends greatly on nutrient recycling processes. Although phytoplankton-zooplankton interactions can contribute a large fraction of nutrient recycling, a simple example reveals recycling by bacteria is also important. A typical value of 0.15 for BP:PP and a typical value of 0.20 for bacterial growth efficiency (BGE, defined as the ratio of BP to bacterial DOC uptake) show DOC uptake by bacteria could be 75% of PP. By using a DOC:DON molar ratio of 15 (as observed at the Hawaii Ocean Time-series program) and C:N for PP of 6.7 (Redfield 1958), the DON recycled by bacteria would be 34% of the nitrogen requirement for primary production. Considering some components of DOM are refractory with higher C:N ratios (discussed below), the actual C:N ratio of DOM taken up by bacteria should be even lower than 15, which will lead to a higher estimate of the bacteria-induced nitrogen recycling.

Bacteria can also take up dissolved inorganic nutrients (DIN). More than half of bacterial nitrogen requirements can be supported by inorganic forms in many circumstances (Kirchman 2000). Thus bacteria can compete with phytoplankton for nutrients and therefore have potential influence on phytoplankton community structure and production.

Based on these considerations, some microbial ecologists and oceanographers have argued to represent the microbial loop explicitly in marine biogeochemical models (e.g., Ducklow 1994, Kirchman 2004, Hood et al. 2006). However, despite the ecological significance of microbial loops in oceanic systems, there are few such models. Even most

of the recent large-scale biogeochemical models do not explicitly include bacteria as a component (e.g., Oschlies & Garcon 1999, Christian et al. 2002, Moore et al. 2002, Gregg et al. 2003, Hood et al. 2003). In these models, simple parameterizations are generally applied for dissolved and particulate organic matter remineralization. Nevertheless, there are several exceptions. Anderson and Pondaven (2003) and Mongin et al. (2003) developed one-dimensional biogeochemical models for the Sargasso Sea that included an explicit bacterial component as well as DOM pools. Laws et al. (2000) constructed a model with simple but explicit bacteria and DOM dynamics. In addition, there are some models designed specifically to study microbial loops (e.g., Anderson & Williams 1999, Anderson & Ducklow 2001). Yet the fact that most biogeochemical models ignore microbial processes is understandable. The biogeochemistry of bacteria and DOM is complex and our current understanding is limited. However, I believe our current understanding of the microbial loop is sufficient to construct an explicit microbial component that is compatible with the current complexity of most biogeochemical models. In addition, we can also use modeling as a tool to test poorly understood microbial mechanisms, e.g., as done by Anderson and Ducklow (2001) to constrain BGE and BP:PP values.

The marine DOM pool represents a broad continuum in terms of its biological reactivity (Carlson 2002), from refractory material turning over on time scales of centuries to labile material turning over on time scales of minutes to days, and semi-labile material with turnover times of months to a few years. That means bacteria do not use all DOM at equal rates. With spatial and temporal variations in DOM composition, there is no uniform rate value to apply to either all oceans or all seasons. At least labile and

semi-labile DOM need to be considered in models, as both Anderson and Pondaven (2003) and Mongin et al. (2003) did. But those two models used constant parameter values for rates of labile and semi-labile DOM usage, which is not consistent with the continuum of DOM reactivity and may decrease the quality of model performance. To deal with DOM production, these models partitioned each DOM source into labile and semi-labile pools at a fixed ratio. By doing this along with fixed DOM reactivity, the necessity of using two DOM pools might be questioned. If the dynamics of DOM reactivity and partitioning could be introduced into a new model, significant improvements could be expected.

When considering whether bacteria take up or regenerate DIN, most of the models discussed above used the formulation proposed by Goldman et al. (1987), which can be represented as (using N as an example):

$$E_N = C_T \left(N:C\right)_S - BGE \cdot C_T \left(N:C\right)_B , \quad (1)$$

where E_N is the regeneration/uptake rate of inorganic N by bacteria, C_T is the total DOC uptake by bacteria, $\left(N:C\right)_S$ is the N/C ratio of substrate (i.e., DOM), $\left(N:C\right)_B$ is the fixed N/C ratio of bacterial biomass. Thus the first term on the right represents the total DON uptake. In the second term on the right, $BGE \cdot C_T$ represents the net C assimilation after respiration, thus this term represents the total N requirement based on assimilated C. Therefore, this equation calculates the difference between the DON uptake and the actually required N. A positive E_N means that the DON uptake exceeds the requirement and bacteria regenerate inorganic N. A negative E_N means the DON uptake

is not sufficient and bacteria have to take up inorganic N to satisfy their N requirement. By varying regeneration/uptake of inorganic N, bacteria can maintain a fixed cellular N/C ratio despite varying substrate composition. This equation indicates that bacteria are more likely to regenerate inorganic N from DOM with a higher N/C and a lower BGE; and to take up inorganic N with an opposite condition.

Although this model is plausible, it has problems when taking the observations into account. Kirchman (2000) used a typical open ocean BGE value of 0.15 and $(C:N)_B$ of 5.9, so bacteria would only take up inorganic N when $(C:N)_S$ is larger than 29. The global mean DOC:DON of 16 suggests bacteria generally excrete inorganic N unless BGE is extremely high and $C:N_B$ is extremely low. Thus in most open ocean conditions, this model will predict inorganic nitrogen regeneration by bacteria. However, observations show bacteria often take up ammonium and/or nitrate (summarized in Kirchman 2000), which indicates the model may not be appropriate under most conditions.

There are several possible explanations why bacteria take up DIN while the model of Goldman et al. (1987) predicts regeneration. First, del Giorgio and Cole (1998) summarized published *in situ* BGE and BP data and found BGE may increase with BP. That is reasonable if considering metabolic energy requirements by bacteria such that under low C supply most of the gross production has to support basal respiration. Second, a flexible stoichiometry of bacteria cells may allow taking extra nutrients even when their carbon component is relatively low. Third, Tupas et al. (1994) observed bacteria assimilated and excreted ammonium simultaneously and also used DON at the same time. Simultaneous assimilation of DOM and inorganic N may be more realistic than letting DIN act as a compensating nutrient. Although this approach may not change the ambient

DIN concentration because bacteria excrete and assimilate simultaneously, bacterial production may be limited by available DIN even when that element in DOM is relatively high. Finally, DIN may not be the only form in which bacteria excrete their surplus elements. Kirchman (2000) argued bacteria may excrete DON compounds other than amino acids that may escape the detection in most experiments.

Spitz et al. (2001) developed a marine ecosystem model including a microbial loop scheme which also considered simultaneous DIN uptake and excretion. However, the fundamental assumption of the scheme may be unrealistic. It assumed that DOM can be divided into two pools: a nitrogenous DOM pool with a constant C:N ratio and a carbohydrate pool with the extra carbon. First, the uptake of nitrogenous DOM including DON and DOC was only determined by the DON concentration, and the carbohydrate uptake was also only determined by the carbohydrate concentration. Therefore the total DOC uptake was only determined by its concentration. Second, this extra carbohydrate pool enhanced ammonium uptake into the bacteria, but inhibited the ammonium excretion from the bacteria. In the other words, a higher C:N ratio of total DOM pool, thus more carbohydrate, was assumed to enhance the N assimilation of bacteria. But as we discussed above a higher C:N ratio of DOM probably indicates a more refractory lability and both DOC and DON should be taken up at lower rates. Bacteria are unlikely to take more ammonium when the DOM is more refractory. The microbial loop scheme of Spitz et al. (2001) ignored the lability of DOM and perhaps even implicitly represented it in a wrong direction.

With the introduction of genomic technology, scientist now have identified tremendous groups of *Bacteria* and *Archaea* and their community structure varies

temporally and spatially (Giovannoni & Rappé 2000, Fuhrman & Hagström 2008). The coexisting of these groups indicates bacteria must functions differently and take different niches, as what we have seen in macro ecosystems. It will be an ultimate goal to represent multiple functional groups of heterotrophic bacteria in a marine ecosystem model. However, our current understandings are limited. We do not have enough observations to set up and parameterize several bacterial groups in one model. For instance, what are their different requirements for DOM? How their BGE varies? Are they subject to different predation? Without the basic information, a good start is to set up only one heterotrophic bacterial group and focus the effort on how it interacts with bulk labile and semi-labile DOM pools. Observed bacterial production and DOM concentration are available to constrain the model by comparing to simulated temporal variations.

Nitrogen Fixation

In a nitrogen-limited ocean, the new or exported production is generally supported by vertical nitrate flux from the deep ocean and N_2 fixation at the surface. When nitrate is vertically transported to the surface, dissolved inorganic carbon will also be transported with it. Thus the exported production supported by the vertical nitrate flux is unlikely to reduce the CO_2 inventory in the surface ocean significantly. Unlike vertical nitrate flux, N_2 fixation is independent of dissolved carbon fluxes from the deep ocean and therefore can potentially drive a net uptake of atmospheric CO_2 and export of carbon. In terms of the carbon cycle and air-sea carbon balance, N_2 fixation may be more important than vertical nitrate flux. Although there are still significant uncertainties and debates about global N_2 fixation rates, a consensus has been reached that a significant fraction of the nitrogen requirement for new production in the tropics and subtropics comes from N_2

fixation, and therefore it supports a significant fraction of the carbon export to the deep ocean (e.g., Carpenter & Romans 1991, Michaels et al. 1996, Capone et al. 1997, Gruber & Sarmiento 1997, Karl et al. 1997, Hood et al. 2002, Casciotti et al. 2008).

Based on studies of *Trichodesmium* spp., a dominant N_2-fixing species in the ocean (Carpenter & Capone 1992, Capone et al. 1997), the factors favoring N_2 fixation have been proposed to be high temperature (Carpenter & Capone 1992, Capone et al. 1997, Subramaniam et al. 2001), strong stratification or high light (Hood et al. 2001, Moore et al. 2002), sufficient iron supply (Raven 1988, Hood et al. 2000, Moore et al. 2004) and oligotrophic conditions with sufficient phosphorus supply (Hood et al. 2000, Dyhrman et al. 2002).

Most reported *Trichodesmium* blooms occur in subtropical and tropical marine habitats with surface water temperature higher than 25ºC (Carpenter & Capone 1992). One possible reason is that low temperature would inhibit the activity of the N_2-fixation enzyme, nitrogenase (Karl et al. 2002). However, because temperature and nitrate are significantly negatively correlated in the marine environment, it is also possible that the selection against N_2-fixing bacteria occurs under conditions of high ambient nitrate where temperature is low.

N_2 fixation demands a significant amount of cellular energy including the direct costs mostly for splitting the bond of the N_2 molecule, as well as the energy requirements to maintain an appropriate cellular environment such as low oxygen, trace-element incorporation, resynthesis of enzymes, etc (Karl et al. 2002). Strong stratification helps to maintain *Trichodesmium* in surface water and obtain high radiation energy, and therefore is proposed as a necessary condition to foster N_2 fixation.

The oligotrophic conditions here generally refer to low concentration of ambient bioavailable nitrogen such as NO_3^-, NH_4^+, urea and some forms of dissolved organic N. Because N_2 fixation requires more energy than direct uptake of fixed N, selection against N_2 fixation will happen under high ambient fixed-N concentration, as has been confirmed by some laboratory studies (Ohki et al. 1991, Mulholland et al. 1999). N_2 fixation is mostly active in oligotrophic oceans. However, the selection of N_2 fixation has to require excess bioavailable P over N because P is also a necessary element for the growth of diazotrophs (Karl et al. 1997). The possible denitrification may reduce the bioavailable N at certain ocean regions (Gruber & Sarmiento 1997, Deutsch et al. 2001). Other studies suggested additional P sources such as atmospheric deposition, the passive upward flux of low density, P-enriched organic matter (Karl & Tien 1997, Karl et al. 2002) and the upwelling of low nitrate:phosphate seawater (Karl & Letelier 2008). In addition, it also has been proposed that *Trichodesmium* are able to sink below the nutricline and take up phosphorus, and then return to the surface with enhanced buoyancy from carbohydrate ballasting. This migratory P acquisition relaxes the phosphorus limitation on N_2-fixation (Karl et al. 1998, Clayton 2001).

Iron is a critical metal co-factor for nitrogenase (Howard & Rees 1996), and the growth of *Trichodesmium* based on N_2 fixation could cost ~ 5 times more iron than the growth based on ammonium (Kustka et al. 2003). Iron is generally depleted in the surface waters (Johnson et al. 1997) and has been shown to limit *Trichodesmium* N_2 fixation under field conditions (Rueter et al. 1992, Paerl et al. 1994).

These conclusions are generally derived from observations or theoretical assumptions about *Trichodesmium* spp. and therefore may not be applicable for recently

discovered unicellular diazotrophic cyanobacteria such as croccosphaera (Falcón et al. 2004, Church et al. 2005). Estimates from both the Pacific and the Atlantic suggested N_2 fixation by unicellular N_2-fixers could be comparable to that by *Trichodesmium* spp. (Capone 2001, Zehr et al. 2001, Carpenter et al. 2004, Falcón et al. 2004, Montoya et al. 2004). Unfortunately, we almost know nothing about unicellular N_2-fixers. In comparison with *Trichodesmium* spp. that mainly exist as colonies, it can be expected that unicellular fixers may be subject to more intensive grazing pressure. They also appear to be tolerant of mixing events because some observations show they occur in deeper water below 100 m (Montoya et al. 2004). The iron requirement for unicellular N_2-fixers is still uncertain. However, it appears that both *Trichodesmium* spp. and unicellular N_2-fixers are favored by oligotrophic conditions, when in competition with other phytoplankton.

If these hypotheses are true, climate change caused by human activities can influence or even trigger oceanic N_2 fixation (Boyd & Doney 2002). For example, a warmer climate may increase water temperature and produce strong stratification in the ocean to support the development of N_2 fixation. Climate change may also change iron supply by influencing wind patterns and dust deposition (Moore et al. 2006). However, the relationship between climate change and N_2 fixation seems complicated and there is still no direct evidence to confirm these ideas.

Some of the current models regarding N_2-fixation have represented diazotrophs explicitly. For example, Hood et al. (2001) modeled *Trichodesmium* biomass explicitly, and this model was coupled to a three-dimensional circulation model of the tropical and subtropical Atlantic (Coles et al. 2004, Hood et al. 2004). The N_2-fixation rates in these models depend on light and therefore mixed layer depth. But these applications did not

include iron or phosphorus limitation.

In the model by Tyrrell (1999), the N_2-fixation rate depends on the supply ratio of inorganic nitrogen and phosphorus. In the model by Neumann (2000) the N_2-fixation rate depends on phosphorus supply, temperature and light. These two models distinguish nitrogen and phosphorus control of phytoplankton and diazotroph growth, but a fixed Redfield N:P ratio is assumed for the organic matter.

Fennel et al. (2002) developed a model with N_2-fixation and both N and P limitation that also allows for different stoichiometric N:P ratios in different organic matter compartments and exports. But this model did not provide a mechanism for enhancing phosphorus supply to the surface by vertical migration of N_2-fixers. Moore et al. (2002)'s model included diazotrophs in a multi-species, multi-element biogeochemical model with nitrogen, phosphorus and iron limitation, and flexible stoichiometric ratios in all model compartments. Although this model still has no mechanism of enhancing phosphorus supply, the flexible N:P ratios and the application of nutrient restoring terms in the nutricline appears to alleviate the phosphorus supply problem.

Flexible C:N:P Stoichiometry

Since Redfield (1934) observed the elemental composition of phytoplankton is relatively stable in the world ocean (Redfield ratio, molar ratio of C:N:P = 106:16:1), marine biogeochemical models have tended to use one of the elements as currency (usually N) by assuming other elements are proportional to this currency at the Redfield ratio. This method dramatically simplified the models and led to some success. Although the Redfield ratio seems appropriate to represent the average status of phytoplankton in the ocean (Falkowski 2000, Michaels 2001), some caution should be taken when

applying it to a model. The Redfield ratio perhaps is a general average rather than a specific requirement for phytoplankton growth (Arrigo 2005). It only represents the condition of bulk organic matter, but not each specific biological species or chemical substrate. DOM appears to be depleted in N and P and with large temporal and spatial variability (Benner 2002), which would be one of our major concerns for bacterial production. Even for phytoplankton, the C:N or C:P ratios tend to be higher and photosynthesis does not cease when corresponding nutrients are limited (Geider et al. 1998, Bertilsson et al. 2003). This condition also determines the rate and composition of DOM production by phytoplankton (Fogg 1966, Bjørnsen 1988, Nagata 2000) as phytoplankton may release more DOM when photosynthesis is more nutrient limited. In addition, N_2-fixers tend to require less P when fixing N_2 (Letelier & Karl 1998). All these observations require that a flexible stoichiometry has to be included for modeling heterotrophic bacteria-DOM interactions and N_2 fixation.

Some recent biogeochemical models have given up a fixed C:N:P ratio of phytoplankton and allow flexible elemental ratios. This improvement recognizes biological adaptation to environmental conditions and potentially makes modeling results more consistent with observation. Flexible elemental ratios for phytoplankton were represented in the models of Anderson and Pondaven (2003) and Mongin et al. (2003). The elemental ratios of bacteria in these models were fixed, although the observed C:N ratio of bacteria cells can vary 3 fold (Fukuda et al. 1998) and the C:P ratio varies even more (Tezuka 1990). The variable products of phytoplankton and/or zooplankton excretion are major sources of DOM and could affect the elemental composition of DOM. This in turn determines whether bacteria take up or regenerate DIN as well as the rate of

this uptake/regeneration (Goldman et al. 1987, Kirchman 2000). The variable elemental composition of bacteria will also influence elemental composition of higher trophic levels and in turn modify regeneration of DIN by zooplankton. Bacterial regulation of DIN will therefore influence the model simulation of primary production.

Data Assimilation

There are more and more observations becoming available for oceanographers. For example, two long-term ocean observatories, the Hawaii Ocean Time-series (HOT) and Bermuda Atlantic Time-Series Study (BATS), have conducted monthly marine biogeochemical observations since 1988 (Karl et al. 2001, Doney & Ducklow 2006). How to combine these numerous data with marine ecosystem models becomes a major challenge for ocean scientists. One of the solutions is to conduct data assimilation. Data assimilation has been widely used in meteorology and ocean circulation modeling to improve forecasts from numerical models or to improve the numerical model itself. More recently, it has been used with ecosystem models (e.g. Fasham & Evans 1995, Lawson et al. 1996, Spitz et al. 1998, Vallino 2000, Spitz et al. 2001, Fennel et al. 2002, Friedrichs et al. 2007).

Ecosystem modeling results are determined by model structure, parameter values, physical forcings and initial and boundary conditions. If a model has been constructed using up-to-date ecological knowledge, and because physical forcings, initial and boundary conditions are directly estimated from observations, the major uncertainties are contributed by the model parameters. Many model parameters are not measured directly, and large spatial and temporal variability exists for many other parameters. Data assimilation can be used to minimize the differences between modeling results and

observations by objectively optimizing model parameters.

The perspectives of data assimilation can help in the following ways. (1) As field measurements are expensive and/or difficult, only a small number of the state variables and rates are directly measured. Data assimilation can use the observations to help improve confidence in other unknown variables and rates. (2) Probably the main limitation of direct ocean observations is that the study sites are remote and difficult to access and therefore the time resolution of data is very coarse. An optimized model may fill the gaps between observations. (3) Assimilating data for different years/seasons can provide temporal variability of optimized parameters and the system status, which may be useful to study the mechanisms controlling the ecosystem. (4) Assimilating data for different models and comparing optimized modeling results can provide useful information to identify robustness of different models and important model structures. (5) Assimilating different data and looking at the confidence of the resulting optimized parameters can identify the important data types which may be useful to guide future observations.

The variational adjoint method (Long & Thacker 1989a, b, Lawson et al. 1995) is one of the commonly used data assimilation methods to optimize simulations towards observations by varying model parameters. The method requires 4 components: (1) a forward model supported by physical forcings and initial and boundary conditions; (2) a cost function evaluating misfits between the forward model results and the assimilated observations; (3) an adjoint of the forward model to compute the gradient of the cost function with respect to model parameters; and (4) an optimization procedure which uses the computed gradient from the adjoint model to determine the direction and the optimal

step size by which the parameters need to be modified in order to reduce the cost function. The 4 components are used iteratively to adjust parameter values until preset criteria are met (e.g., low gradient).

Spitz et al. (2001) did a data assimilation for a marine ecosystem model using the BATS time-series. However, the study did not include the bacterial biomass and production rates in the assimilation process, although the model has a microbial loop scheme. In addition, the study did not consider the temporal variability of optimal parameter values with only about 5 years of data being assimilated. My study will assimilate 15 years of data and can be expected to illuminate some patterns of temporal variability.

Study Site

The long-term biogeochemical observatory, Station ALOHA (22.75ºN, 158.00ºW), was established by the Hawaii Ocean Time-series (HOT) project in the North Pacific Subtropical Gyre (NPSG) in 1988 (Figure 1). The station is 100 km north of Oahu, Hawaii with a bottom depth of 4700 m. Because the station is within the weak southward return flow of the eastern gyre and the eddies around the station are not energetic (Karl et al. 2001), horizontal transport is not significant. Therefore the station is a good site to apply 1-D ecological modeling.

The station has a relatively stable environment. It is quite stratified at the surface with an observed average mixed layer depth of 67 m. Strong seasonal deep mixing is unusual with winter mixed layer only reaching ca. 120 m, and nutrient upwelling is

limited. It is a typical oligotrophic environment with average nitrate concentration below 0.05 μM and phosphate below 0.1μM in the upper 50 m, indicating a nitrogen-limited system. The water temperature is warm and does not show strong seasonal variations: on average 23.9^{o} C, 25.7^{o} C, 26.1^{o} C and 24.0^{o} C at the surface for the four seasons. However, seasonal variability of primary production is obvious with high rates during summer as indicated by the HOT time-series. Thus the station is potentially a good test site for model robustness on how to capture seasonal biological variations in this physically stable environment. In addition, this station satisfies our central study issues for bacteria – DOM dynamics and N_2 fixation processes. As the stratification persists, semi-labile DOM may stay in the surface over seasons to years, which provides us a test of DOM dynamics with the model. The station has a low f-ratio, i.e., primary production is mainly supported by recycled nutrients, and a corresponding low ratio of export to primary production (ca. 27 mg C/m^2/day vs. 500 mg C/m^2/day). It can be expected that the heterotrophic bacteria may be more important at ALOHA than at other high f-ratio sites in recycling nutrients and supporting primary production. The warm, stratified and oligotrophic condition is likely preferred for N_2 fixation. Actual N_2-fixing *Trichodemium* spp. blooms have been observed frequently around Station ALOHA area (e.g. Letelier & Karl 1996, Karl et al. 1997, Subramaniam et al. 2001, Westberry & Siegel 2006, Fong et al. 2008).

Of course it is the long-term observations that Station ALOHA will benefit our modeling study most. Since it was established in 1988, roughly monthly measurements have been conducted, and numerous data have become available for the major physical, biological and chemical variables (Karl et al. 2001, Doney & Ducklow 2006) including

those directly facilitating the setup of the model: CTD based physical and irradiance measurements, nutrients, primary production, phytoplankton and heterotrophic bacteria cell counts, mesozooplankton, dissolved organic matter, particulate organic matter, sediment traps, etc. The long-term data are a good source to calibrate, validate and improve the model, as well as to conduct studies of temporal variability. The data sets are publicly available online at http://hahana.soest.hawaii.edu/hot/hot-dogs/interface.html.

I will also use two supplementary sites from the Arabian Sea and Equatorial Pacific to compare the differences among them and Station ALOHA. Because seasonal monsoons occur in the Arabian Sea and strong upwelling occurs across the Equatorial Pacific compared to the strongly stratified condition at the Station ALOHA, these two sites may show different mechanisms for our central issues of microbial loops and N_2 fixation. I use observations from the U.S. Joint Global Ocean Flux Study (JGOFS) seasonal cruises during late 1994 – 1995 for the Arabian Sea and 1992 for the Equatorial Pacific. Station S7 (16ºN, 62ºE) was selected for the Arabian Sea and the equatorial station (ca. 0ºN, 140ºW) was selected for the Equatorial Pacific. As a testbed framework for these two sites has already been set up by others (Friedrichs et al. 2006, Friedrichs et al. 2007), the comparison can be conducted by plugging my ecosystem model into the testbed, along with some recent improvements. Because the testbed at the Arabian Sea and Equatorial Pacific only covers about one year, I will use 2002 of Station ALOHA for the comparison because the data coverage of bacterial production is best for this year.

Scientific Questions

This dissertation will answer following scientific questions:

(1) What controls heterotrophic bacterial production? How do heterotrophic bacteria influence uuper-ocean carbon cycle and nutrient regeneration? Are heterotrophic bacterial dynamics at Hawaii different from the other study sites in the Arabian Sea and Equatorial Pacific?

(2) Is semi-labile DOM an important resource for heterotrophic bacterial production? Does its accumulation in the surface ocean over seasons and years influence primary production and other components of the ecosystem?

(3) How much does N_2 fixation support new primary production at Station ALOHA? What are the differences between *Trichodesmium* spp. and unicellular N_2-fixing cyanobacteria in their dynamics and influences on the system?

(4) What is the value of long-term observations? Did the ecosystem at Station ALOHA change during the past decade, especially primary production, microbial loops and N_2 fixation? What mechanism generates the shift, if it exists?

(5) What model parameters can be constrained by the current available observations? What parameters are important but still not constrained by observations? Specifically, are the heterotrophic bacterial production measurements useful to constrain unknown parameters? What missing measurements should be recommended to include in new observations?

Dissertation Structure

To answer the above questions, this dissertation will conduct 1-dimensinal (1-D) ecosystem modeling and data assimilation focusing on Station ALOHA.

Chapter 1 (this chapter) introduces the scientific questions.

Chapter 2 describes the methods. It first develops an ecosystem model including the structure and the background information to justify the structure. State variables, model equations and parameters are listed. The second part describes the setup of the data assimilative framework for Station ALOHA. It explains how the model grid, initial and boundary conditions and physical forcings were defined. Next are brief descriptions of how to evaluate the model-data misfits through a weighted cost function; how to set up adjoint code to compute the gradients of cost function versus model parameters; and how the limited-memory quasi-Newton method works to optimize the parameters. Finally, the chapter gives the method of setting up the default values of the model parameters.

Chapter 3 conducts system modeling for 2002 at Station ALOHA and uses data assimilation to optimize parameters to fit the observations of that year. Then food web analysis is conducted to track average carbon, nitrogen and phosphorus stocks and fluxes. The analysis focuses on heterotrophic bacteria and dissolved organic carbon flux, primary production by N_2 fixers, nutrient regeneration by heterotrophic bacteria versus zooplankton and carbon export. Possible explanations for these features are proposed. The results at Station ALOHA are compared with results from the Arabian Sea and Equatorial Pacific using the same ecosystem model. The differences are explained with possible reasons.

Chapter 4 focuses on the temporal variability of the system at Station ALOHA. Data assimilation is conducted the whole period of 1991 - 2005. Therefore an optimal parameter set is obtained to represent the best solution for the whole period. The temporal variability is studied by analyzing the average food web structure in each year. Primary

production, dynamics of heterotrophic bacteria and DOM and N_2 fixation are the major issues. The mechanisms controlling the temporal variability are studied to explore the influence of both physical forcings and biogeochemistry (e.g., DOM inventory and composition) on how state changes can accumulate in the system and influence future periods.

Chapter 5 studies how heterotrophic bacterial production (BP) data can constrain the system through data assimilation. This question is evaluated by comparing paired simulations assimilating / not assimilating BP data in 2000 – 2005 when 13 BP profiles were measured at Station ALOHA. Without assimilating BP rates, the bacterial respiration (BR) and therefore the bacterial growth efficiency cannot be constrained. Varying the BR rate within the literature supported range shows a significant influence on the ecosystem. Thus it is suggested to routinely measure BP at Station ALOHA. In addition, other unconstrained, but model-sensitive parameters, are listed to suggest the key measurements to be done in the future.

Chapter 6 summarizes the results from Chapter 3 – 5 and concludes on those scientific questions introduced in Chapter 1.

References

Anderson TR, Ducklow HW (2001) Microbial loop carbon cycling in ocean environments studied using a simple steady state model. Aquat. Microb. Ecol. 26:37-49

Anderson TR, Pondaven P (2003) Non-redfield carbon and nitrogen cycling in the Sargasso Sea: pelagic imbalances and export flux. Deep-Sea Res. I 50:573-591

Anderson TR, Williams PJl (1999) A one-dimensional model of dissolved organic carbon cycling in the water column incorporating combined biological-photochemical decomposition. Global Biogeo. Cycles 13:337-349

Arrigo KR (2005) Marine microorganisms and global nutrient cycles. Nature 437:349-355

Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad F (1983) The ecological role of water-column microbes in the sea. Mar. Ecol. Prog. Ser. 10:257-263

Benner R (2002) Chemical composition and reactivity. In: Hansell DA, Carlson CA (eds) Biogeochemistry of marine dissolved organic matter. Academic Press, San Diego, p 59-90

Berelson WM (2001) The Flux Of Particulate Organic Carbon Into The Ocean Interior: A Comparison Of Four US JGOFS Regional Studies. Oceanography 14:59-67

Bertilsson S, Berglund O, Karl DM, Chisholm SW (2003) Elemental composition of marine Prochlorococcus and Synechococcus: Implications for the ecological stoichiometry of the sea. Limnol. Oceanogr. 48:1721-1731

Bjørnsen PK (1988) Phytoplankton exudation of organic matter: Why do healthy cells do

it? Limnol. Oceanogr. 33:151-154

Boyd PW, Doney SC (2002) Modelling regional responses by marine pelagic ecosystems to global climate change. Geophys. Res. Lett. 29:1806, Doi: 10.1029/2001gl014130

Buesseler KO, Antia AN, Chen M, Fowler SW, Gardner WD, Gustafsson O, Harada K, Michaels AF, van der Loeff'o MR, Sarin M, Steinberg DK, Trull T (2007) An assessment of the use of sediment traps for estimating upper ocean particle fluxes. J. Mar. Res. 65:345-416

Capone DG (2001) Marine nitrogen fixation: what's the fuss? Current Opinion in Microbiology 4:341-348

Capone DG, Zehr JP, Paerl HW, Bergman B, Carpenter EJ (1997) Trichodesmium, a globally significant cyanobacterium. Science 276:1221-1229

Carlson CA (2001) Chapter 15: Production, Consumption and Cycling of Dissolved Organic Matter in the Ocean. In: Hansell DA, Carlson CA (eds) Biogeochemistry of Marine Dissolved Organic Matter. Academic Press, New York

Carlson CA (2002) Production and removal processes. In: Hansell DA, Carlson CA (eds) Biogeochemistry of marine dissolved organic matter. Elsevier Science, USA, New York, p 91-151

Carlson CA, Ducklow HW, Michaels AF (1994) Annual flux of dissolved organic carbon from the euphotic zone in the northwestern Sargasso Sea. Nature 371:405-408

Carpenter EJ, Capone DG (1992) Significance of *Trichodesmium* blooms in the marine nitrogen cycle. In: Carpenter EJ, Capone DG, Rueter JG (eds) Marine Pelagic Cyanobacteria: Trichodesmium and other Diazotrophs. Kluwer Academic

Publishers, Dordrecht, p 211-217

Carpenter EJ, Romans K (1991) Major role of the cyanobacterium Trichodesmium in nutrient cycling in the North Atlantic Ocean. Science 254:1356-1358

Carpenter EJ, Subramaniam A, Capone DG (2004) Biomass and primary productivity of the cyanobacterium, *Trichodesmium* spp., in the tropical N Atlantic Ocean. Deep Sea Research Part I 51:173-203

Casciotti KL, Trull TW, Glover DM, Davies D (2008) Constraints on nitrogen cycling at the subtropical North Pacific Station ALOHA from isotopic measurements of nitrate and particulate nitrogen. Deep Sea Research Part II: Topical Studies in Oceanography 55:1661-1672

Christian JR, Verschell MA, Murtugudde R, Busalacchi AJ, McClain CR (2002) Biogeochemical modelling of the tropical Pacific Ocean. I: Seasonal and interannual variability. Deep Sea Research Part II: Topical Studies in Oceanography 49:509-543

Church MJ, Jenkins BD, Karl DM, Zehr JP (2005) Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the oligotrophic North Pacific Ocean. Aquat. Microb. Ecol. 38:3-14

Clayton TD (2001) Ph.D. dissertation: Trichodesmium spp.: numerical studies of resource competition, carbohydrate ballasting, and remote-sensing reflectance. Ph.D. dissertation, Old Dominion University

Coles VJ, Hood RR, Pascual M, Capone DG (2004) Modeling the effects of Trichodesmium and nitrogen fixation in the Atlantic Ocean. J. Geophys. Res. 109:C06007, doi:10.1029/2002JC001754

del Giorgio PA, Cole JJ (1998) Bacterial Growth Efficiency in Natural Aquatic Systems. Ann. Rev. Ecol. Syst. 29:503-541

Deutsch C, Gruber N, Key RM, Sarmiento JL (2001) Denitrification and N_2 fixation in the Pacific Ocean. Global Biogeochem Cycles 15:483-506

Doney SC, Ducklow HW (2006) A decade of synthesis and modeling in the US Joint Global Ocean Flux Study. Deep-Sea Research Part Ii-Topical Studies in Oceanography 53:451-458

Ducklow HW (1994) Modeling the microbial foodweb. Microb. Ecol. 28:303-319

Ducklow HW (2000) Bacterioplankton production and biomass in the oceans. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley, New York, p 85-120

Ducklow HW, Carlson CA (1992) Oceanic bacterial productivity. Adv. Microb. Ecol. 12:113-181

Ducklow HW, Carlson CA, Bates NR, Knap AH, Michaels AF (1995) Dissolved organic carbon as a component of the biological pump in the North Atlantic Ocean. Philosophical Transactions of the Royal Society of London B 348:161-167

Ducklow HW, Steinberg DK, Buesseler KO (2001) Upper Ocean Carbon Export and the Biological Pump. Oceanography 14:50-58

Dyhrman ST, Webb EA, Anderson DM, Moffett JW, Waterbury JB (2002) Cell-specific detection of phosphorus stress in Trichodesmium from the Western North Atlantic. Limnol. Oceanogr. 47:1832-1836

Emerson S, Quay P, Karl D, Winn C, Tupas L, Landry M (1997) Experimental determination of the organic carbon flux from open-ocean surface waters. Nature 389:951-954

Eppley RW, Peterson BJ (1979) Particulate organic matter flux and planktonic new production in the deep ocean. Nature 282:677-680

Falcón LI, Carpenter EJ, Cipriano F, Bergman B, Capone DG (2004) N_2 fixation by unicellular bacterioplankton from the Atlantic and Pacific oceans: Phylogeny and in situ rates. Appl. Envir. Microbiol. 70:765-770

Falkowski PG (2000) Rationalizing elemental ratios in unicellular algae. Journal of Phycology 36:3-6

Fasham MJR, Evans GT (1995) The Use of Optimization Techniques to Model Marine Ecosystem Dynamics at the Jgofs Station at 47°N 20°W. Philosophical Transactions of the Royal Society of London Series B 348:203-209

Fennel K, Spitz YH, Letelier RM, Abbott MR, Karl DM (2002) A deterministic model for N_2 fixation at stn. ALOHA in the subtropical North Pacific Ocean. Deep-Sea Res. II 49:149-174

Fogg GE (1966) The extracellular products of algae. Oceanogr. Mar. Biol. Ann. Rev. 4:195-212

Fong AA, Karl DM, Lukas R, Letelier RM, Zehr JP, Church MJ (2008) Nitrogen fixation in an anticyclonic eddy in the oligotrophic North Pacific Ocean. Isme Journal 2:663-676

Friedrichs MAM, Dusenberry JA, Anderson LA, Armstrong RA, Chai F, Christian JR, Doney SC, Dunne J, Fujii M, Hood R, McGillicuddy DJ, Moore JK, Schartau M, Spitz YH, Wiggert JD (2007) Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. Journal of Geophysical Research-Oceans 112:C08001, Doi 10.1029/2006jc003852

Friedrichs MAM, Hood RR, Wiggert JD (2006) Ecosystem model complexity versus physical forcing: Quantification of their relative impact with assimilated Arabian Sea data. Deep-Sea Res. II 53:576-600

Fuhrman JA, Hagström Å (2008) Bacterial and Archael community structure and its patterns. In: Kirchman DL (ed) Microbial Ecology of the Oceans, 2nd edition. John Wiley & Sons, Inc., Hoboken, New Jersey, p 45-90

Fukuda R, Ogawa H, Nagata T, Koike I (1998) Direct determination of carbon and nitrogen contents of natural bacterial assemblages in marine environments. Appl. Envir. Microbiol. 64:3352-3358

Geider RJ, MacIntyre HL, Kana TM (1998) A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. Limnol. Oceanogr. 43:679-694

Giovannoni SJ, Rappé M (2000) Evolution, diversity and molecular ecology of marine prokaryotes. In: Kirchman DL (ed) Microbial Ecology of the Oceans, 1st edition. Wiley-Liss, New York, p 47-84

Goldman JC, Caron DA, Dennett MR (1987) Regulation of gross growth efficiency and ammonium regeneration in bacteria by substrate C:N ratio. Limnol. Oceanogr. 32:1239-1252

Gregg WW, Ginoux P, Schopf PS, Casey NW (2003) Phytoplankton and iron: validation of a global three-dimensional ocean biogeochemical model. Deep-Sea Res. II 50:3143-3169

Gruber N, Sarmiento JL (1997) Global patterns of marine nitrogen fixation and denitrification. Global Biogeo. Cycles 11:235-266

Hansell DA (2002) DOC in the global ocean carbon cycle. In: Hansell DA, Carlson CA (eds) Biogeochemistry of Marine Dissolved Organic Matter. Academic Press, San Diego, p 685-715

Hedges JI (1992) Global biogeochemical cycles: progress and problems. Mar. Chem. 39:67-93

Honjo S, Manganini SJ, Krishfield RA, 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 76:217-285

Hood DW, Michaels AF, Capone DG (2000) Answers sought to the enigma of marine nitrogen fixation. EOS 81:133, 138-139

Hood RR, Bates NR, Capone DG, Olson DB (2001) Modeling the effect of nitrogen fixation on carbon and nitrogen fluxes at BATS. Deep Sea Research Part II: Topical Studies in Oceanography 48:1609-1648

Hood RR, Coles VJ, Capone DG (2004) Modeling the distribution of *Trichodesmium* and nitrogen fixation in the Atlantic Ocean. J. Geophys. Res. 109:C06006, doi:10.1029/2002JC001753

Hood RR, Kohler KE, McCreary JP, Smith SL (2003) A four-dimensional validation of a coupled physical-biological model of the Arabian Sea. Deep-Sea Res. II 50:2917-2945

Hood RR, Laws EA, Armstrong RA, Bates NR, Brown CW, Carlson CA, Chai F, Doney SC, Falkowski PG, Feely RA, Friedrichs MAM, Landry MR, Moore JK, Nelson DM, Richardson TL, Salihoglu B, Schartau M, Toole DA, Wiggert JD (2006)

Pelagic functional group modeling: Progress, challenges and prospects. Deep-Sea Res. II 53:459-512

Hood RR, Subramaniam A, May LR, Carpenter EJ, Capone DG (2002) Remote estimation of nitrogen fixation by Trichodesmium. Deep-Sea Res. II 49:123-147

Howard JB, Rees DC (1996) Structural basis of biological nitrogen fixation. Chem. Rev. 96:2965-2982

Johnson KS, Gordon RM, Coale KH (1997) What controls dissolved iron concentrations in the world ocean? Authors' closing comments. Mar. Chem. 57:181-186

Karl D, Michaels A, Bergman B, Capone D, Carpenter E, Letelier R, Lipschitz F, Paerl H, Sigman D, Stal L (2002) Dinitrogen fixation in the world's oceans. Biogeochemistry 57/58:47-98

Karl DM, Dore JE, Lukas R, Michaels AF, Bates NR, Knap A (2001) Building the long-term picture: the U.S. JGOFS time-series programs. Oceanography 14:6-17

Karl DM, Hebel DV, Bjoerkman K, Letelier RM (1998) The role of dissolved organic matter release in the productivity of the oligotrophic North Pacific Ocean. Limnol. Oceanogr. 43:1270-1286

Karl DM, Letelier R, Tupas L, Dore J, Christian J, Hebel D (1997) The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. Nature 388:533-538

Karl DM, Letelier RM (2008) Nitrogen fixation-enhanced carbon sequestration in low nitrate, low chlorophyll seascapes. Marine Ecology-Progress Series 364:257-268

Karl DM, Tien G (1997) Temporal variability in dissolved phosphorus concentrations in the subtropical North Pacific Ocean. Mar. Chem. 56:77-96

Kirchman DL (2000) Uptake and regeneration of inorganic nutrients by marine heterotrophic bacteria. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, Inc., New York, p 261-288

Kirchman DL (2004) A primer on dissolved organic material and heterotrophic prokaryotes in the oceans. In: Follows M, Oguz T (eds) The Ocean Carbon Cycle and Climate. Kluwer Academic Publishers, Dordrecht, p 31-64

Kustka AB, Sanudo-Wilhelmy SA, Carpenter EJ, Capone DG, Burns J, Sunda WG (2003) Iron requirements for dinitrogen- and ammonium-supported growth in cultures of Trichodesmium (IMS 101): Comparsion with nitrogen fixation rates and iron:carbon ratios of field populations. Limnol Oceanogr 48:1869-1884

Landry M (2001) Microbial loops. In: Steele JH, Thrope SA, Turekian KK (eds) Encyclopedia of Ocean Sciences. Academic Press, London, p 1763-1770

Laws EA, Falkowski PG, Smith WO, Jr., Ducklow HW, McCarthy JJ (2000) Temperature effects on export production in the open ocean. Global Biogeo. Cycles 14:1231-1246

Lawson LM, Hofmann EE, Spitz YH (1996) Time series sampling and data assimilation in a simple marine ecosystem model. Deep Sea Research Part II: Topical Studies in Oceanography 43:625-651

Lawson LM, Spitz YH, Hofmann EE, Long RB (1995) A data assimilation technique applied to a predator-prey model. Bulletin of Mathematical Biology 57:593-617

Letelier RM, Karl DM (1996) Role of Trichodesmium spp. in the productivity of the subtropical North Pacific Ocean. Mar. Ecol. Prog. Ser. 133:263-273

Letelier RM, Karl DM (1998) Trichodesmium spp. physiology and nutrient fluxes in the

North Pacific subtropical gyre. Aquat. Microb. Ecol. 15:265-276

Long RB, Thacker WC (1989a) Data Assimilation into a Numerical Equatorial Ocean Model .1. The Model and the Assimilation Algorithm. Dynamics of Atmospheres and Oceans 13:379-412

Long RB, Thacker WC (1989b) Data Assimilation into a Numerical Equatorial Ocean Model .2. Assimilation Experiments. Dynamics of Atmospheres and Oceans 13:413-439

Michaels AF (2001) Elemental stoichiometry, new production and nitrogen fixation. Oceanography 14:68-74

Michaels AF, Olson D, Sarmiento JL, Ammerman JW, Fanning K, Jahnke R, Knap AH, Lipschultz F, Prospero JM (1996) Inputs, losses and transformations of nitrogen and phosphorus in the pelagic North Atlantic Ocean. Biogeochemistry 35:181-226

Mongin M, Nelson DM, Pondaven P, Brzezinski MA, Treguer P (2003) Simulation of upper-ocean biogeochemistry with a flexible-composition phytoplankton model: C, N and Si cycling in the western Sargasso Sea. Deep-Sea Res. I 50:1445-1480

Montoya JP, Houll CM, Zehr JP, Hansen A, Villareal TA, Capone DG (2004) High rates of N_2 fixation by unicellular diazotrophs in the oligotrophic Pacific Ocean. Nature 430:1027-1032

Moore JK, Doney SC, Kleypas JA, Glover DM, Fung IY (2002) An intermediate complexity marine ecosystem model for the global domain. Deep-Sea Res. II 49:403-462

Moore JK, Doney SC, Lindsay K (2004) Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. Global Biogeochem Cycles

18:GB4028, doi:10.1029/2004GB002220

Moore JK, Doney SC, Lindsay K, Mahowald N, Michaels AF (2006) Nitrogen fixation amplifies the ocean biogeochemical response to decadal timescale variations in mineral dust deposition. Tellus Series B 58:560-572

Mulholland MR, Ohki K, Capone DG (1999) Nitrogen utilization and metabolism relative to patterns of N2 fixation in cultures of Trichodesmium NIBB1067. Journal of Phycology 355:977-988

Nagata T (2000) Production mechanisms of dissolved organic matter. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, New York, p 121-152

Neumann T (2000) Towards a 3D-ecosystem model of the Baltic Sea. J. Mar. Syst. 25:405-419

Ohki K, Zehr JP, Falkowski PG, Fujita Y (1991) Regulation of nitrogen-fixation by different nitrogen sources in the marine non-heterocystous cyanobacterium Trichodesmium sp. NIBB1067. Arch. Microbiol. 156:335-337

Olson JS, Garrels RM, Berner RA, Armentano TV, Dyer MI, Taalon DH (1985) The natural carbon cycle. In: Trabalka JR (ed) Atmospheric Carbon Dioxide and the Global Carbon Cycle. US Department of Energy, Washington, DC, p 175-213

Oschlies A, Garcon V (1999) An eddy-permitting coupled physical-biological model of the North Atlantic - 1. Sensitivity to advection numerics and mixed layer physics. Global Biogeo. Cycles 13:135-160

Paerl HW, Prufert-Bebout L, Guo C (1994) Iron-stimulated N_2 fixation and growth in natural and cultured populations of the planktonic marine cyanobacterium *Trichodesmium* sp. Appl. Envir. Microbiol. 60:1044-1047

Raven JA (1988) Limits to growth. In: Borowitzka MA, Borowitzka LJ (eds) Micro-algal Biotechnology. Cambridge University Press, Cambridge, p 331-356

Redfield AC (1934) On the proportions of organic derivations in sea water and their relation to the composition of plankton. In: Daniel RJ (ed) James Johnstone Memorial Volume. University Press of Liverpool, p 177-192

Redfield AC (1958) The biological control of chemical factors in the environment. American Scientist 46:205-221

Rueter JG, Hutchins DA, Smith RW, Unsworth NL (1992) Iron nutrition of *Trichodesmium*. In: Carpenter EJ, Capone DG, Rueter JG (eds) Marine Pelagic Cyanobacteria: *Trichodesmium* and other Diazotrophs. Kluwer Academic Publishers, The Netherlands, p 289-306

Spitz YH, Moisan JR, Abbott MR (2001) Configuring an ecosystem model using data from the Bermuda Atlantic Time Series (BATS). Deep Sea Research Part II: Topical Studies in Oceanography 48:1733-1768

Spitz YH, Moisan JR, Abbott MR, Richman JG (1998) Data assimilation and a pelagic ecosystem model: parameterization using time series observations. J. Mar. Syst. 16:51-68

Subramaniam A, Brown CW, Hood RR, Carpenter EJ, Capone DG (2001) Detecting Trichodesmium blooms in SeaWiFS imagery. Deep-Sea Res. II 49:107-121

Sverdrup HU, Johnson MW, Fleming RH (1946) The oceans, their physics, chemistry, and general biology. Prentice Hall, New York

Tezuka Y (1990) Bacterial Regeneration of Ammonium and Phosphate as Affected by the Carbon - Nitrogen - Phosphorus Ratio of Organic Substrates. Microb. Ecol.

19:227-238

Tupas LM, Koike I, Karl DM, Holm-Hansen O (1994) Nitrogen metabolism by heterotrophic bacterial assemblages in Antarctic coastal waters. Polar Biology 14:195-204

Tyrrell T (1999) The relative influences of nitrogen and phosphorus on oceanic primary production. Nature 400:525-531

Vallino JJ (2000) Improving marine ecosystem models: use of data assimilation and mesocosm experiments. J. Mar. Res. 58:117-164

Westberry TK, Siegel DA (2006) Spatial and temporal distribution of Trichodesmium blooms in the world's oceans. Global Biogeo. Cycles 20:GB5016

Williams PJl (2000) Heterotrophic bacteria and the dynamics of dissolved organic material. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, New York, p 153- 200

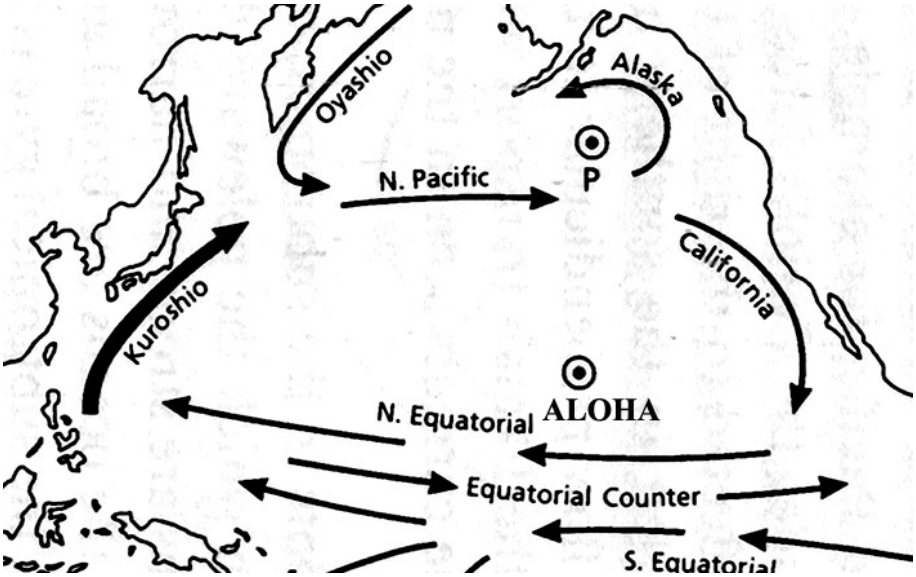
Williams PM, Druffel ERM (1987) Radiocarbon in dissolved organic carbon in the central north Pacific Ocean. Nature 330:246-248

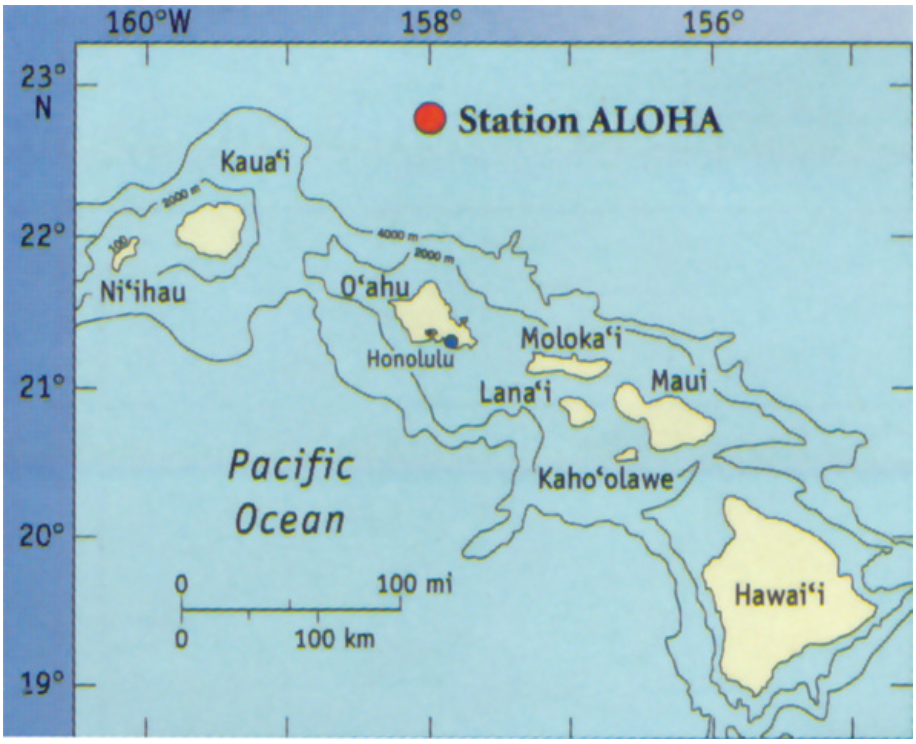
Zehr JP, Waterbury JB, Turner PJ, Montoya JP, Omoregie E, Steward GF, Hansen A, Karl DM (2001) Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. Nature 412:635-638

Figure Caption

Figure 1. The location of the study site: Station ALOHA of Hawaii Ocean Time-series (HOT) project. Upper panel, circulation pattern of North Pacific Subtrophic Gyre (NPSG) and the location of Station ALOHA; Bottom panel, detailed deployment areas of Station ALOHA.

Figure 1


Oyashio
Alaska
N. Pacific
P
California
Kuroshio
N. Equatorial
ALOHA
Equatorial Counter
S. Equatorial


160°W
158°
156°
23° N
22°
21°
20°
19°
Station ALOHA
Kaua'i
Ni'ihau
O'ahu
Honolulu
Moloka'i
Lana'i
Maui
Kaho'olawe
Hawai'i
Pacific Ocean
0
100 mi
0
100 km

Chapter 2. A Marine Ecosystem Model with Microbial Dynamics and Its Application to a Data Assimilative Testbed for Open Ocean

Abstract

The microbial schemes in most marine biogeochemical models are crude and do not represent current understanding of the microbial dynamics. We constructed a new model to include a more realistic description of microbial dynamics including heterotrophic and diazotrophic bacteria. The model simulates the carbon (C), nitrogen (N) and phosphorous (P) components in phytoplankton, heterotrophic bacteria, protozoan and metazoan zooplankton, labile and semi-labile dissolved organic matter (DOM), as well as chlorophyll *a* and dissolved inorganic nutrients including ammonium, nitrate and phosphate. The C, N and P are tracked independently so that the stoichiometry of the compartments is flexible. The model represents heterotrophic bacterial processes differently from previous models mainly in four ways including: 1) continuous lability of semi-labile DOM determined by nutrient contents, 2) selective semi-labile DOM uptake by bacteria, 3) influence of inorganic nutrients on bacterial growth, and 4) variable bacterial growth efficiency depending on bacterial production. The model represents phytoplankton in 3 compartments: non-diazotrophic picophytoplankton, diazotrophic cyanobacteria of *Trichodesmium* spp. and unicellular coccoid N_2-fixers. N_2 fixation rates are limited by light, temperature and ambient phosphate. Also, high ammonium or nitrate concentration may also inhibit the N_2 fixation rates because in the model the diazotrophs tend to use bio-available inorganic nitrogen before they conduct N_2 fixation. The differences between the two diazotrophs are their dependencies on light and the grazing pressures on them. The model is plugged into an ecological data assimilative testbed for the Arabian Sea, Equatorial Pacific and Hawaii Ocean Time-series (HOT) Station ALOHA. It provides a 1-dimensional (1-D) physical environment for the surface 150 m

(200 m for Station ALOHA), as well as boundary and initial conditions for state variables. It also includes a data assimilative framework that uses the variational adjoint method to minimize the difference between the modeling results and up to 17 types of observations by optimizing the model parameters. The observations applied include nitrate, phosphate, phytoplankton biomass, chlorophyll *a*, primary production, heterotrophic bacterial biomass and production, dissolved organic matter, particulate organic matter and sediment trap data. This objective method can improve the modeling results and constrain some unknown parameter values, and represents one way to apply large data sets to modeling studies.

1. Introduction

Dissolved organic matter (DOM) is the dominant organic pool in the ocean. DOM is processed mainly by heterotrophic bacteria before it can be remineralized into inorganic nutrients. Thus the interaction between DOM and heterotrophic bacteria is one of the major processes to control the carbon cycle and nutrients recycling in the ocean. N_2 fixation conducted by marine cyanobacteria has been proposed as an important new source of fixed N, which significantly influences the new production and export in the world's oligotrophic ocean (e.g. Carpenter & Romans 1991, Gruber & Sarmiento 1997, Karl et al. 1997, Karl et al. 2002, Falcón et al. 2004). A significant fraction of the nitrogen requirement for new primary production in the subtropical oceans comes from N_2 fixation, and therefore it supports a significant fraction of the carbon export to the deep ocean (e.g., Carpenter & Romans 1991, Michaels et al. 1996, Capone et al. 1997, Gruber & Sarmiento 1997, Karl et al. 1997, Hood et al. 2002). Both *Trichodesmium* spp. and unicellular N_2-fixing cyanobacteria were reported to be the important N_2 fixers in the global ocean (Carpenter & Capone 1992, Capone et al. 1997, Zehr et al. 2001, Karl et al. 2002, Montoya et al. 2004). A marine ecosystem model with detailed microbial dynamics was developed to address these issues. The model has flexible stoichiometry, which provides necessary frame to realize the microbial dynamics and also improves the representation of the other schemes of the system.

The long-term biogeochemical observatory, Station ALOHA, was established by the Hawaii Ocean Time-series (HOT) project in the North Pacific Subtropical Gyre (NPSG) in 1988. Since then roughly monthly measurements have been conducted and numerous data have become available for the major physical, biological and chemical variables.

Because the station is within the weak southward return flow of the eastern gyre and the patterns of eddies around the station are not energetic (Karl et al. 2001), the horizontal transport is not significant. Therefore the station is a good site to apply 1-D ecological modeling. The long-term data are a good source to calibrate, validate and improve the model. The data sets are publicly available online at http://hahana.soest.hawaii.edu/hot/hot-dogs/interface.html.

Two supplementary sites from the Arabian Sea and the Equatorial Pacific were used to compare the differences among them and Station ALOHA. Because seasonal monsoons occur in the Arabian Sea and strong upwelling occurs across the Equatorial Pacific compared to the strongly stratified condition at the Station ALOHA, these two sites may show different mechanisms for our central issues of microbial loops and N_2 fixation. I used observations from the U.S. Joint Global Ocean Flux Study (JGOFS) seasonal cruises during late 1994 – 1995 for the Arabian Sea and 1992 for the Equatorial Pacific. The data sets are publicly available online at http://usjgofs.whoi.edu/.

To embed the observations with the model, we constructed a data assimilative framework to minimize the difference between the modeling results and the observations by objectively optimizing parameter values through a variational adjoint method. It is expected to constrain the model parameters, improve the modeling results, indicate potential mechanisms driving temporal variability, and appreciate the importance of different observations.

In this chapter, I describe in Section 2 a marine ecosystem model with flexible stoichiometry and intermediate complexity of microbial schemes including microbial loop and N_2 fixation. The model is applied to a data assimilative testbed set up for Station

ALOHA, the Arabian Sea and the Equatorial Pacific, which will be discussed in Section 3.

2. Ecosystem Model Description

The structure and the state variables of the model are shown in Figure 1 and Table 1 and will be discussed below. The model tracks carbon (C), nitrogen (N) and phosphorus (P) cycles through state variables except for inorganic nutrients. The model also tracks chlorophyll *a* (Chl *a*) for 3 phytoplankton variables. The full parameters and equations of the model are listed in Table 2 and the Appendix A.

2.1. State Variables

The names and definitions of state variables are listed in Table 1. Phytoplankton (PHY) in the model represents both eukaryotic and prokaryotic autotrophs. Only picophytoplankton are considered in order to represent the dominant forms in oligotrophic open ocean gyres. There are two forms of diazotrophic cyanobacteria represented in the model: *Trichodesmium* spp. (TR) and unicellular N_2-fixing cyanobacteria (UN) which conduct photosynthesis and fix N_2 when ambient ammonium and nitrate levels are low. Bacteria (BA) represent heterotrophic prokaryotes including both *Bacteria* and *Archaea* using DOM as their sole food source. Protozoa (PRT) are protozoan zooplankton capable of ingesting PHY, UN and BA. Metazoa (MZ) are metazoan zooplankton that graze on TR and PRT. DOM is divided into labile (LDOM) and semi-labile (SDOM) pools. Refractory DOM is not explicitly represented because its turnover time is much longer than the other two pools and would not significantly

influence the system. However, fluxes to refractory DOM are implicitly represented as loss terms for some state variables. Detritus (DET) in the model is not divided into different sizes but represents an average pool. All these state variables are independently tracked as C, N and P, and additional Chl *a* for PHY, TR and UN. Nutrients in the model include 2 forms of inorganic N, ammonium (NH4) and nitrate (NO3), and 1 form of inorganic P, phosphate (PO4).

2.2. Parameters

Table 2 lists all the model parameters and their default values. Most parameters are "adjustable" that is, subject to optimization by data assimilation. Their default values were obtained by conducting data assimilation (see Section 3 for details). Another set is some parameters of UN, which are assigned equal values to other parameters of SP and TR, i.e., they change with other parameters. The 3rd set is "quota" parameters, which are cellular molar N:C and P:C ratios of PHY, TR, UN, BA, PRT and MZ defined as "common", "minimum" and/or "maximum" quota for each state variable, in which Redfield ratios are used as "common" quota for phytoplankton variables. They are assigned fixed values and are not adjustable during the optimization of the model.

2.3. Temperature Effect

Growth, respiration and N_2 fixation rates in the model are affected by water temperature. The values of these parameters (Table 1) are defined at 25ºC and Arrhenius function is adopted to increase/decrease them with increasing/decreasing temperature (Appendix A 1.1). The Arrhenius parameter of 4000 used here (Moore et al. 2002)

increases the rates by 57% if the temperature increases 10ºC to 35 ºC, or decreases the rates by 38% if the temperature decreases 10ºC to 15ºC. These rates include maximum growth rates of PHY, TR, UN, BA, PRT and MZ, maximum N_2 fixation rates of TR and UN, basal respiration rates of BA, PRT and MZ, turnover rate of LDOM and maximum turnover rate of SDOM. In addition, a temperature below 15^{o} C will completely cease the N_2 fixation of TR. Because active respiration rates are related to growth rates, which have already been adjusted by temperature, active respiration rates are not changed independently by the temperature effect.

2.4. Phytoplankton (PHY) Processes

We adopted the scheme of Geider et al. (1998) (GD98 hereafter) to simulate phytoplankton photosynthesis, nitrogen consumption and the synthesis of Chl *a* with some modifications. The nitrogen consumption scheme of GD 98 is also applied to phosphorus assimilation although it may not be perfect because phosphorus can be "luxury consumed" in great excess of current needs (Armstrong 2006). However, my model focuses more on nitrogen cycle than phosphorus cycle. And the problem is also partially alleviated by given a higher maximum phosphorus quota for phytoplankton as discussed below.

As proposed by Droop (1974), the phytoplankton carbon growth rate is limited by the cellular nutrient contents or "cell quota". We follow the modified form of GD98 where the nutrient to carbon biomass ratios, i.e., N:C and P:C, replace the cell quota in Droop (1974). Thus the N-limited term I used is slightly different from GD98:

$$Nfun_{PHY} = \frac{Q^C_{N,PHY} - q^C_{N,\min,PHY}}{q^C_{N,rdf,PHY} - q^C_{N,\min,PHY}}, \text{ then set } 0 \le Nfun_{PHY} \le 1, \quad (1)$$

where $Q^C_{N,PHY}$ is the cellular molar N:C ratios of PHY, and $q^C_{N,\min,PHY}$ and $q^C_{N,rdf,PHY}$ are parameters of the minimum and common allowable cellular molar N:C ratios. The P-limited term $Pfun_{PHY}$ has the same form as $Nfun_{PHY}$ by using quota parameters for phosphorus. So that the light-saturated photosynthesis rate is:

$$P^C_{\max,PHY} = \mu_{PHY} \cdot Tfun \cdot \min\left(Nfun_{PHY}, Pfun_{PHY}\right), \quad (2)$$

where μ_{PHY} is the parameter for maximum C-specific growth rate. The Redfield ratio is used for q_{rdf} in the model because some recent studies showed it is robust as a mean value for phytoplankton composition (Falkowski 2000, Michaels 2001). Therefore Equation (1) and (2) represents that the photosynthesis rate starts to drop when the cell quota becomes lower than the Redfield ratio until it is completely shutdown when the cell quota reaches the minimum value $q_{\min}$.

The actual photosynthesis rate will also be limited by light intensity as:

$$GROW^C_{PHY} = C_{PHY} \cdot P^C_{\max,PHY} \cdot \left(1 - EXP\left(\frac{-\alpha^{CHL}_{PHY} \cdot Q^C_{CHL,PHY} \cdot PAR}{P^C_{\max,PHY}}\right)\right) \cdot EXP\left(-\beta_{PHY} \cdot PAR\right), \quad (3)$$

where C_{PHY} is the carbon biomass of PHY, α^{CHL}_{PHY} is the parameter for initial slop of photosynthesis versus PAR (photosynthetically active radiation) curve, $Q^C_{CHL,PHY}$ is the ratio of Chl *a* to carbon biomass, β_{PHY} is the light-inhibition parameter. This equation is slightly different from GD98 with a light inhibition term $EXP\left(-\beta_{PHY} \cdot PAR\right)$

included.

For N and P consumption, an opposite limiting term is used (using N as an example):

$$V^{N}_{\max,PHY} = \frac{q^{C}_{N,\max,PHY} - Q^{C}_{N,PHY}}{q^{C}_{N,\max,PHY} - q^{C}_{N,rdf,PHY}}, \text{ then set } 0 \le V^{N}_{\max,PHY} \le 1, \quad (4)$$

where $q^{C}_{N,\max,PHY}$ is the parameter of maximum allowable cellular molar N:C ratio. Thus the N or P consumption is not limited by cellular quota if it is lower than the Redfield ratio, but starts to drop when the quota is higher than the Redfield ratio until the consumption is totally shutdown when the quota reaches $q_{\max}$. The consumption is also limited by the ambient nutrient level with a Monod function and the actual consumption rates for ammonium and nitrate is:

$$GROW^{NH4}_{PHY} = C_{PHY} \cdot v^{N}_{REF,PHY} \cdot Tfun \cdot V^{N}_{\max,PHY} \cdot \frac{NH4}{NH4 + k^{NH4}_{PHY} + NO3 \cdot k^{NH4}_{PHY} / k^{NO3}_{PHY}}, \quad (5)$$

$$GROW^{NO3}_{PHY} = C_{PHY} \cdot v^{N}_{REF,PHY} \cdot Tfun \cdot V^{N}_{\max,PHY} \cdot \frac{NO3}{NO3 + k^{NO3}_{PHY} + NH4 \cdot k^{NO3}_{PHY} / k^{NH4}_{PHY}}, \quad (6)$$

where $v^{N}_{REF,PHY}$ is the maximum N consumption rate per C biomass, k^{NH4}_{PHY} and k^{NO3}_{PHY} is the half-saturation concentration of ammonium and nitrate. This form avoids the total N consumption rate to exceed $v^{N}_{REF,PHY}$ and represents a ammonium inhibition on nitrate consumption with a lower k^{NH4}_{PHY} than k^{NO3}_{PHY}. The phosphate consumption is the similar Monod function as the N consumption of Equation (5) and (6) but without an inhibition term in the denominator. The model also considers the required respiration in

order to reduce the consumed nitrate (Appendix A2.6), but ignores the energy requirement for ammonium and phosphate consumption.

It is difficult to determine the minimum and maximum quotas. Goldman et al. (1979) obtained molar N:C between 0.05 - 0.14 and molar P:C between 0.001 – 0.03 under different nutrient and light levels. Falkowski et al. (1985) observed molar N:C between 0.11 - 0.23 for three phytoplankton species under different light levels. GD98 used 0.047 and 0.23 (0.19 for one studying species) for minimum and maximum molar N:C of the four phytoplankton species in their model. In my model, $q^C_{N,\min,PHY}$ and $q^C_{N,\max,PHY}$ are set to be same as GD98, but $q^C_{P,\min,PHY}$ and $q^C_{P,\max,PHY}$ are set to be $\pm 80\%$ from $q^C_{P,rdf,PHY}$ (Table 2). These parameters roughly represent the observed pattern that P:C varies more than N:C and are within the observed ranges, which also allows an alleviated limitation on P consumption.

A scenario of primary production as described by the model could be: (1) At the beginning of day, the N:C cellular ratio of phytoplankton tends to be high as the phytoplankton has consumed N during last night. Thus the C assimilation rate will be high while the N assimilation is low at early of the day, which will reduce cellular N:C ratio and therefore gradually reduce the photosynthesis rate. If the N:C ratio reaches $q_{\min}$, the C assimilation is totally shutdown. (2) During night, the C assimilatoin is totally shutdown; the low N:C resulted during day makes N consumption high and cellular N content will be restoring. The increasing N:C will slow down the N assimilation rate. If N:C reaches $q_{\max}$, N consumption will be completely shutdown.

The Chl *a* production rate is 0 if PAR = 0, and otherwise is determined as:

$$GROW_{PHY}^{CHL} = \theta \cdot GROW_{PHY}^{N} \cdot \frac{GROW_{PHY}^{C}}{\alpha_{PHY}^{CHL} \cdot CHL_{PHY} \cdot PAR \cdot EXP\left(-\beta_{PHY} \cdot PAR\right)}, \quad (7)$$

where θ is the maximum Chl *a* to N ratio. The last multiplier on the right represents the ratio of energy assimilated to energy absorbed as described by GD98. The equation represents a variable allocation of assimilated N to Chl *a*, and a low PAR level will generate a high Chl production.

The model considers the DOM production of phytoplankton in 3 path ways (Appendix A 2.7), passive release of labile DOM, active release of carbohydrate and other DOM.

Two models have been proposed to explain extracellular release of DOM by phytoplankton. The passive diffusion model (Fogg 1966, Bjørnsen 1988) considers that low-molecular-weight (LMW) DOM will passively cross the phytoplankton cell membrane because of the concentration gradient from cell to environment. In our model, the passively released LMW DOM has the same C:N:P ratio as phytoplankton biomass (Biddanda & Benner 1997) and its exudation rate is assumed to be 5% of biomass per day (Bjørnsen, 1988). We assume these LMW materials such as neutral sugars and dissolved free amino acids (DFAA) are most labile and therefore go directly to the labile DOM pool.

Another model, the overflow model (Fogg 1966, Nagata 2000), assumes that carbon production by photosynthesis is relatively independent of inorganic nutrient incorporation and therefore fixed carbon could be more than the amount that can be supported by the incorporated nutrients. Under this condition, some DOM with high carbon to nutrient

ratio may be actively released (excreted). However, this assumption is not always valid as the high-molecular-weight (HMW) DOM production is commonly observed despite of the nutritional conditions. Carbohydrates is one of the major forms of the released DOM (Biddanda & Benner 1997, Biersmith & Benner 1998) and the model assumes 5% of gross primary production is released as carbohydrate. These carbohydrate can be labile (Amon & Benner 1994, Santschi et al. 1995), but can also seasonally accumulate in the water (Carlson et al. 1994, Carlson et al. 1998). The model roughly divided the released carbohydrate as 75% of labile dissolved organic carbon (DOC) and 25% of semi-labile DOC.

The model also considers that phytoplankton release DOM in order to adjust their stoichiometry to approach the Redfield ratio. We assume this type of release only happens when cellular organic C is in excess, and it does not occur when cellular organic C is low. This is reasonable because nutrients are generally limiting in the ocean and this mechanism helps phytoplankton to compete with other organisms by storing nutrients when ambient concentrations are high. This function calculates the amount of excess C relative to the most limiting nutrient (N or P) in phytoplankton biomass, and assumes this amount of C is released on a time scale of 2 days. Also, under the excess carbon condition, there would be another excess element, either N or P, when comparing these two element quotas. Bronk (2002) argued the overflow model is unlikely to impact dissolved organic nitrogen (DON) release significantly because most environments that are lacking in P generally do not have an excess of DON. However, phytoplankton also release DON and dissolved organic phosphorous (DOP) (Nagata 2000, Carlson 2002) and there is no evidence that these DON and DOP are exclusively from passive release. In

other words, DON or DOP may also be actively released by phytoplankton. To address these two issues, a lower release rate as 1/4 that of the DOC, i.e., a time scale of 8 days, is used when DON or DOP are released.

Morán et al. (2001) found that the phytoplankton DOC production (release) increases with the particulate organic carbon (POC) production, but the ratio of DOC to POC production has negative relation to POC production. The easiest model to realize this observation is a constant release term that independent with the primary production, and a variable term that is proportional to the primary production, as the scheme I used here for a passive release and a active carbohydrate release.

We also assume phytoplankton can aggregate and form sinking particles via a concentration-dependent process that is proportional to the square of the biomass (Appendix A 2.8). The phytoplankton removal by protozoa grazing is discussed in the section on zooplankton processes.

Chl *a* is removed with aggregation and grazing processes of phytoplankton, but is not influenced by the DOM release.

2.5. Diazotrophic Phytoplankton Processes

There are two different diazotrophic compartments represented in the model, *Trichodesmium* spp. (TR) and Unicellular N_2-fixers (UN). The equations controlling their C growth rate are the same as for PHY, although different parameter values are used to set light limitation and cell quotas. A much smaller initial slope of production versus light (α_{TR}^{CHL}) puts a stronger light limitation for TR and limits the majority of its production to the surface ocean (Letelier & Karl 1996, 1998). The maximum growth rate and the initial

slope of UN (α_{UN}^{CHL}) are smaller than PHY, but the initial slope is large enough to allow UN to exist at depths down to the base of the euphotic zone (Zehr et al. 2001, Montoya et al. 2004). The cellular N:C quota of TR parameters are set to be slightly higher than PHY to represent the observations for *Trichodesmium* spp. (Letelier & Karl 1996) and to slightly stimulate N2 fixation. The P:C quota parameters of TR are set to be significantly smaller than PHY to represent lower P requirements by *Trichodesmium* spp. as observed by some studies (Letelier & Karl 1996, Fu et al. 2005). As there is no data for the stoichiometry of unicellular diazotrophs, UN uses the same quota parameter values as TR.

TR and UN take up dissolved inorganic nutrients similar to PHY in the model (Carpenter & McCarthy 1975, Ohki et al. 1991, Mulholland et al. 1999), but with higher TR half-saturation concentration for inorganic nutrients than PHY. Most TR cells in the ocean aggregate as colonies with larger size than single cells, and they are expected, therefore, to take up inorganic nutrients at lower rates. We assume UN has the same half-saturation concentrations for inorganic nutrient uptake as PHY. However, certainly the major difference of N_2-fixers from PHY is their ability to fix N_2 to alleviate the limitation by dissolved inorganic nitrogen (DIN). To simulate N_2 fixation, the model first calculate the maximum total N requirement:

$$GROW_{\max,TR}^{N} = C_{TR} \cdot v_{REF,TR}^{N} \cdot Tfun \cdot V_{\max,TR}^{N}, \quad (8)$$

where $v_{REF,TR}^{N}$ is the maximum N growth rate per C biomass, $V_{\max,TR}^{N}$ is the cell quota limitation with the same form as PHY (Equation 4). The model first fulfills this requirement by DIN (ammonium and nitrate) consumption because they use less energy than N_2 fixation (Karl et al. 2002). If the DIN consumption is not enough to satisfy N

requirement, the model triggers N_2 fixation.

To calculate N_2 fixation rate, the model then assumes under optimal condition, the ratio of total N to total C biomass after production and respiration equals to the maximum cellular N:C quota $q^C_{N,\max,PHY}$ (using TR as an example) :

$$\frac{N_{TR} + GROW^{NO3}_{TR} + GROW^{NH4}_{TR} + NFIX_{\max,TR}}{C_{TR} + GROW^{C}_{TR} - \zeta^{NO3} \cdot GROW^{NO3}_{TR} - \zeta^{NF} \cdot NFIX_{\max,TR}} = q^C_{N,\max,TR}, \quad (9)$$

where $NFIX_{\max,TR}$ is the maximum N_2 fixation rate of TR, ζ^{NF} is the energy (respiration) requirement rate per N_2 fixation. Together with temperature control and cell quota control, the maximum N_2 fixation rate can be derived from Equation (9) as:

$$NFIX_{\max,TR} = \frac{\left(C_{TR} + GROW^{C}_{TR} - \zeta^{NO3} \cdot GROW^{NO3}_{TR}\right) \cdot q^C_{N,\max,TR} - N_{TR} - GROW^{NO3}_{TR} - GROW^{NH4}_{TR}}{1 + \zeta^{NF} \cdot q^C_{N,\max,TR}} \cdot Tfun \cdot V^N_{\max,TR}. \quad (10)$$

For TR, if temperature drops below 15ºC, $NFIX_{\max,TR}$ is set to 0. The realized N_2 fixation plus the DIN uptake is bounded by the maximum N requirement $GROW^N_{\max,TR}$. Both nitrate consumption and N_2 fixation requires carbon respiration (Appendix A 3.5).

Both TR and UN take up phosphate similarly as PHY. But a possible advantage of *Trichodesmium* spp. was proposed by some studies that when phosphate is depleted at the surface they can control their buoyancy to sink during night to take up phosphate and then buoys back to surface (Letelier & Karl 1998, Villareal & Carpenter 2003, White et al. 2006). Our model implicitly represents this process by proposing that TR can refill and make its P quota to $\left(q^C_{P,\max,TR} + q^C_{P,rdf,TR}\right)/2$ in a time scale of 1 day. Thus this P refill will be another new supply of P to the model.

Trichodesmium spp. passively releases DOM (Capone et al. 1994, Glibert & Bronk 1994) and ammonium (Mulholland & Capone 2001, Mulholland et al. 2004, Mulholland 2007). In addition to the same passive and active release of DOM and carbohydrate as PHY, a portion of N fixed by N_2 fixation will be released as ammonium and DON (Appendix A 3.7). The rate varies in different studies (Glibert & Bronk 1994, Mulholland et al. 2004) and the model assumes 18% of N fixed by N_2 fixation is released as ammonium and another 18% is released as labile DON. Although there is no clear result showing how unicellular N_2-fixers release DOM or ammonium, the model assumes they are similar to TR.

The model-defined competition among PHY, UN and TR allows coexistence, but the dominant species may vary with changing environmental conditions. PHY has a higher light-saturate growth rate, so that it dominates when the ambient nutrients are replete. If the nutrients are depleted, PHY growth is limited and N_2-fixers, either UN or TR, may dominate. Because UN has a higher initial slope of the P vs. E curve, it will dominate over TR in the deeper layer. However, near the surface (e.g. upper 50 m) where phosphate may be depleted, the ability of TR to take phosphate from deep water by buoyancy regulation may help it compete over UN. In addition, when light is not low, TR can do more N_2 fixation than UN because TR has a higher maximum N_2 fixation rate. Another feature of the model is the different grazing rates on TR and UN. TR is mostly in colonies and may be also toxic to zooplankton (Hawser et al. 1992). In the model TR is consumed only by metazoa at a lower rate compared to UN, which is consumed by protozoa at a higher rate.

2.6. Bacterial Processes

The DOM is divided into labile and semi-labile DOM. The model determines the lability of SDOM by its N or P content. There is evidence that the N/C and P/C profiles of DOM decrease from surface to depth (Benner 2002, Ogawa & Tanoue 2003). As we have already known that the DOM with low N/C or P/C ratio in the deep ocean are mostly refractory (Ogawa & Tanoue 2003), it is safe to assume that the lability of SDOM increases with its N:C or P:C ratio and vice versa. Equation (11) is used to establish a continuum of SDOM lability by controlling the amount of semi-labile DOC available during a time unit. However, the lability of LDOM is assumed to be independent on its N:C or P:C ratios and all LDOM is available during a time unit. A basic consideration is that some carbohydrates released by phytoplankton may be labile but without N or P constituents. Therefore it is not appropriate to limit the lability of labile DOM by its N:C or P:C ratio.

The Available Semi-labile DOC (ASC) for bacterial uptake during a time unit is represented as:

$$ASC = C_{SDOM} \cdot \min\left(1,\ EXP\left(b_{LABI,N} \cdot \left(Q^{C}_{N,SDOM} / q^{C}_{N,BA} - 1\right)\right),\ EXP\left(b_{LABI,P} \cdot \left(Q^{C}_{P,SDOM} / q^{C}_{P,BA} - 1\right)\right)\right), \tag{11}$$

where $q^{C}_{N,BA}$ and $q^{C}_{P,BA}$ are bacterial common cellular N:C and P:C and $b_{LABI,N}$ and $b_{LABI,P}$ are unitless adjustable parameters. Theoretically, when both N:C and P:C ratios of semi-labile DOM reach bacterial optimal cellular quota, all semi-labile DOC would be available for bacteria although it will not happen in reality. Reduced elemental ratios in semi-labile DOM would quickly lower its lability.

Bacteria C growth rate depends on both bacterial cellular quota and available labile

and semi-labile DOC. The cellular quota limitation is (using N as an example):

$$Nfun_{BA} = Q^{C}_{N,BA} / q^{C}_{N,BA}\text{, then set } \; Nfun_{BA} \le 1, \tag{12}$$

and a similar form is used to P quota limitation. Therefore bacteria C growth will be limited if its cellular N or P quota is less than the corresponding common quota. The bacterial labile and semi-labile DOC consumption is:

$$GROW^{LDOC}_{BA} = \mu_{BA} \cdot Tfun \cdot C_{BA} \cdot \min\left(Nfun_{BA,} Pfun_{BA}\right) \cdot \frac{ALC}{ALC + k_{DOM} + ASC}, \tag{13}$$

$$GROW^{SDOC}_{BA} = \mu_{BA} \cdot Tfun \cdot C_{BA} \cdot \min\left(Nfun_{BA,} Pfun_{BA}\right) \cdot \frac{ASC}{ASC + k_{DOM} + ALC}, \tag{14}$$

where μ_{BA} is the C-specific maximum bacterial growth rate, ALC is available labile DOC which equals to labile DOC concentration in the model, k_{DOM} is half-saturate concentration of available DOC for bacterial growth. Equation (13) and (14) does not artificially increase the bacterial C growth rate to exceed μ_{BA} while the bacterial uptake of labile and semi-labile DOC will influence each other. The scheme sets ALC and ASC equally in terms of bacterial preference. But ALC equals the concentration of labile DOM, and ASC is set by Equation (11) as a variable portion of semi-labile DOM. Thus Equation (11) ultimately controls the preferential consumption of bacteria on labile and semi-labile DOM. Because labile and semi-labile DOC are the only sources for bacterial C growth in the model, the summary of Equation (13) and (14) set the bacterial carbon demand (BCD), or the total bacterial carbon consumption.

The model assumes bacteria take labile DOM as a bulk substrate so that the ratio of labile DON (DOP) to labile DOC consumption is same as the bulk N(P):C ratio of labile DOM. However, the model assumes that bacteria are capable of selecting semi-labile DOM with higher N and P contents. This assumption is consistent and necessary for the model mechanism that semi-labile DOM with higher N:C or P:C ratio is more labile (Equation 11). The selection-term controlled semi-labile DON is:

$$GROW_{BA}^{SDON} = GROW_{BA}^{SDOC} \cdot \min\left(q_{N,BA}^{C},\ Q_{N,SDOM}^{C} + f_{SLCT,BA} / Nfun_{BA} \cdot \left(q_{N,BA}^{C} - Q_{N,SDOM}^{C}\right)\right) \quad (15)$$

where $f_{SLCT,BA}$ is the parameter of selection strength varying between 0 and 1. Thus the ratio of semi-labile DON to DOC consumption varies between the bulk N:C of semi-labile DOM $Q_{N,SDOM}^{C}$ and the bacterial common cellular quota $q_{N,BA}^{C}$. The ratio will be moved towards $q_{N,BA}^{C}$ with higher selection strength $f_{SLCT,BA}$ and lower limitation of bacteria cellular quota $Nfun_{BA}$ (therefore lower $Q_{N,BA}^{C}$). Same form is used for semi-labile DOP consumption.

Bacteria take up DIN even if the net flux is releasing DIN as discovered by Tupas et al. (1994) (see Chapter 1). We propose bacteria take up DOM as well as dissolved inorganic nutrients simultaneously. However, that does not mean the mass balance model of Goldman et al. (1987) is not applicable to determine the net flux of inorganic nutrients (see Chapter 1). Because the stoichiometry of bacterial cells does not vary substantially (Kirchman 2000), bacteria have to maintain their elemental composition by taking up or releasing inorganic nutrients. Actually for some modeling purposes, simulation only of the net flux for inorganic nutrients is sufficient. Our model simulates both the influx and outflux of inorganic nutrients in order to compare DOM and inorganic nutrient influxes

and their limiting effect on bacteria. But only the net flux of inorganic nutrients is meaningful for other state variables. Some evidence shows that ammonium and phosphate are readily substrate for bacteria uptake (Kirchman 1994). As the reduced forms, the model assumes they have same activity as labile DOM. The ratio of ammonium (phosphate) to labile DON (DOP) consumption equals to the ratio of their concentration before bacterial cellular quota limitation (use ammonium as an example):

$$GROW_{BA}^{NH4} = GROW_{BA}^{LDON} \cdot NH4 / N_{LDOM} / \min\left(1,\ Nfun_{BA}\right), \tag{16}$$

However, the energetic cost to assimilate nitrate are higher than DON (Vallino et al. 1996) as nitrate requires reduction before being assimilated. But bacterial nitrate uptake was commonly observed (Kirchman & Wheeler 1998, Middelburg & Nieuwenhuize 2000, Allen et al. 2002). The model assumes bacterial nitrate consumption occurs only when the bacterial cellular N:C ratio is less than its common ratio. In other words, bacteria only take up nitrate when they are in short of nitrogen. This assumption is reasonable because there is no reason for bacteria to take up nitrate and to respire energy while their cells have enough nitrogen stored. When the condition is met, the equation used for bacterial nitrate consumption is:

$$GROW_{BA}^{NO3} = \min\left(\begin{array}{l} 0.1 \cdot NO3 \cdot \dfrac{1}{\min\left(1,\ Nfun_{BA}\right)} \cdot \dfrac{GROW_{BA}^{LDON} + GROW_{BA}^{SDON}}{N_{LDOM} + N_{SDOM}}, \\ \left(NO3 + NH4\right) \cdot \dfrac{GROW_{BA}^{LDON} + GROW_{BA}^{SDON}}{N_{LDOM} + N_{SDOM}} - GROW_{BA}^{NH4} \end{array}\right). \tag{17}$$

So the equation puts two limitations on the nitrate uptake: (1) nitrate consumption is no more than 10% of N-specific bulk DOM (including labile and semi-labile) consumption rate, which limits the maximum nitrate consumption rate; and (2) the nitrate plus ammonium consumption is no more than N-specific bulk DOM (including labile and

semi-labile) consumption rate, which set the inhibition of ammonium consumption on nitrate consumption.

Bacterial growth efficiency, BGE, equals to the ratio of bacterial net production (BP) to bacterial carbon demand (BCD), BGE = BP/BCD. BCD = BP + BR, where BR is the bacterial respiration. Thus BGE = BP/(BP+BR), or BR = BCD×(1 – BGE). As BCD has been known, the model has to estimate BGE so that BP can be derived. A paper of del Giorgio and Cole (1998) summarized published *in situ* BGE and BP data and found BGE increases with BP. By definition, a higher BP together with a higher BGE indicates a higher BCD. Thus the observations also support a positive relation between BGE and BCD. The model realizes this relationship with negative correlation between bacterial active respiration and BCD. The respiration of bacterial carbon in the model includes basal respiration, active respiration and the required respiration to reduce assimilate nitrate. The function is:

$$RESP_{BA} = \zeta^{NO3} \cdot GROW_{BA}^{NO3} + resp_{BA}^{B} \cdot Tfun \cdot C_{BA} + \left(resp_{\min,BA}^{A} + \left(resp_{\max,BA}^{A} - resp_{\min,BA}^{A} \right) \cdot EXP\left(-b_{RESP} \cdot GROW_{BA}^{C} \right) \right) \cdot GROW_{BA}^{C} \quad , (18)$$

where $resp_{BA}^{B}$ is the basal respiration rate in day^{-1}, $GROW_{BA}^{C}$ is the bacterial carbon demand (BCD) or the bacterial total DOC uptake, $resp_{\min,BA}^{A}$ and $resp_{\max,BA}^{A}$ are the minimum and maximum portions of gross production actively respired with inversed unit to $GROW_{BA}^{C}$, and b_{RESP} is a positive parameter with inversed unit to $GROW_{BA}^{C}$. By using this function, the BGE can increase with BP (see the modeled BGE versus modeled BP in Figure 5, Chapter 3).

Bacteria can be an important source of refractory DOM from their membranes or

can transform labile DOM to refractory DOM (Tanoue et al. 1995, McCarthy et al. 1998, Carlson 2002). Our data assimilation optimizes the refractory DOC release by bacteria around 2% of their carbon uptake, which is inside the range of other studies (1.4 ~ 8.6% as summarized by Carlson 2002). In the deep ocean where DOM is mostly refractory, DON:DOC and DOP:DOC are ~0.05 and ~0.0007 respectively as shown by the HOT data set, which are the ratios we use for bacterial refractory DOM in the model. The refractory DOM is not explicitly represented in the model and is assumed to be a loss from the model domain.

After the above processes, the model adjusts bacterial elemental composition by either 1) excreting ammonium or phosphate if N or P is in excess relative to C, or 2) excreting semi-labile DOM if C is in excess. A similar function as phytoplankton active release is used to control excretion rates. Compared to phytoplankton excretion, higher values are given to semi-labile DOM release rate $ex_{ADJ,BA}$ and inorganic nutrient regeneration rate $remi_{BA}$ (Appendix A5.9) to reflect that bacteria have higher ability to control their stoichiometry. Bacteria are removed by protozoa which will be discussed in zooplankton processes. Also, there is a mortality term that assumes a certain percentage (2% - 5% resulted by data assimilation) of bacteria biomass per day is lost to labile DOM pool due to virus attack.

In summary, our model represents bacterial processes differently from previous models mainly in four ways by including: 1) continuous variable lability of semi-labile DOM determined by nutrient contents, 2) selective semi-labile DOM uptake by bacteria, and 2) variable bacterial growth efficiency depending on bacterial production rate.

2.7. Zooplankton (PRT and MZ) Processes

We use the same form of grazing function for PRT grazing on PHY, UN and BA, and MZ grazing on TR and PRT. The general form of carbon-specific grazed amount of a prey " *pry* " by a predator " *pdt* " is:

$$GRAZ_{pry}^{C} = Tfun \cdot \mu_{pdt} \cdot C_{pdt} \cdot \frac{C_{pry}{}^{2}}{C_{pry}{}^{2} + g_{pry}{}^{2} + \left(C_{pry2} \cdot g_{pry} / g_{pry2}\right)^{2} + \left(C_{pry3} \cdot g_{pry} / g_{pry3}\right)^{2} + \cdots}, \tag{19}$$

where μ_{pdt} is the maximum grazing rate in day^{-1}, and g_{pry} is the grazing parameter depending on the grazer and prey, the subscripts $pry2$ and $pry3$ etc represent other preys grazed by the same predator pdt. Thus Equation (19) represents a density-dependent grazing and also a preferential selection of a predator on different food sources. N and P components are grazed relative to C at the same ratios as those of prey composition.

Zooplankton release a portion of ingested organic matter as DOM both from sloppy feeding and excretion. The model does not separate these two difference sources, but assumes a certain percentage of the total grazed C by PRT and MZ is release as DOC. The model then partitions the release DOC with 75% to labile pool and the other 25% to semi-labile pool. The ratio of released labile DON (DOP) to released labile DOC is same as the N(P):C ratio of zooplankton food source. The semi-labile DON (DOP) release is treated similarly as labile DON (DOP), but the rate can be increased or reduced when the zooplankton cellular N(P):C ratio is above or below the common ratio. (Appendix A 6.2 & 7.2).

Both PRT and MZ respire C from basal respiration, which respires a portion of their

biomass, and active respiration which respires a portion of their grazed C.

After the above DOM excretion and respiration, PRT and MZ adjust their body quota by either releasing semi-labile DOM if C is in excess, or regenerating inorganic nutrients if N or P is in excess (Appendix A 6.4-5 and 7.4-5). The scheme is similarly to BA.

PRT and MZ also produce POM. The C component of this production is assumed to be a fixed percentage of grazed C. To simplify, a constant C:N:P ratio (molar 220:25:1) is defined for POM production so that the N and P components of it can be calculated from the C component. This ratio is higher than the Redfield ratio to indicate our assumption that zooplankton tends to assimilate higher quality food with higher N:C and P:C ratios, but to egest fecal materials with lower N:C and P:C ratios.

MZ (not applied to PRT) also produces refractory DOM with a N:C and P:C ratio similarly as bacteria. As MZ is the highest trophic level represented in the model, it is subject to a density-dependent removal process by implicit higher trophic levels. The removed C, N and P go to 3 pools, semi-labile DOM, POM and inorganic nutrients, with the partition defined by parameters.

2.8. Other Processes (DET, dissolved inorganic nutrients & DOM)

The POM production from the processes discussed above is distributed to detritus pool and a constant vertical sinking speed is assigned to detritus. Detritus is transformed to semi-labile DOM in the model. Compared to other biogeochemical models that assign a remineralization rate for detritus, detritus has to be transformed to DOM before it can be regenerated to inorganic nutrients. The transformation rate of the P component of the

detritus is higher than its N component, which is in turn higher than that of C. This is supported by the HOT data sets where POM in shallow sediment traps has more N and P than that collected at deeper depth.

The model also includes a simple nitrification process which converts a certain percentage of ammonium to nitrate per day. Nitrification in the euphotic zone is argued to be an important process in order to estimate new production correctly in the world ocean (Yool et al. 2007).

Another parameter was assigned to convert SDOM to refractory DOM at a slow rate depending on the stoichiometry of SDOM (Appendix A 10.1). When N:C and P:C are high, SDOM converts to refractory DOM very slowly because N and P enriched SDOM are more likely labile. The conversion rate can be high when the N:C and P:C ratios of SDOM are very low. The C:N:P ratio of the produced refractory DOM is predefined by parameters.

3. Data Assimilative Testbed

Based on a testbed for the Arabian Sea and Equatorial Pacific (Friedrichs et al. 2007), a vertical 1-D data assimilative testbed was constructed for Station ALOHA, including physical forcings and transportation schemes, initial and boundary conditions, an optimization software and assimilative data sets. The testbed uses the variational adjoint method (Lawson et al. 1995) to optimize parameters towards observations. The method requires four components, all of which are supported by the testbed: (1) a forward model supported by physical forcings and initial and boundary conditions; (2) a cost function evaluating misfits between the forward model results and the assimilated observations; (3)

an adjoint of the forward model to compute the gradient of the cost function with respect to model parameters, generated by auto-differential software TAPENADE 2.1; and (4) an optimization procedure (M1QN3 3.1), which uses the computed gradient from the adjoint model to determine the direction and the optimal step size by which the parameters need to be modified in order to reduce the cost function. Thus the four components are used iteratively to adjust parameter values until preset criteria are met (e.g., low gradient). Theoretically, besides the model structure (i.e. equations), model results depend on model parameters as well as initial conditions, boundary conditions and physical conditions. Data assimilation could also incorporate these other factors as control variables but this is more computation-intensive. For our application, we assume these conditions are derived from observations, and the model spin-up can help to minimize the influence of initial conditions. Thus only model parameters were considered to be control variables and modified by data assimilation.

Although the testbed was constructed initially for the purpose of running the ecosystem model described in Section 2, it was also set up to allow other ecosystem models to be plugged in to use same boundary conditions, physical forcings, parameter optimization, etc.

Unless otherwise stated, most data used to develop the testbed are from the HOT time-series which is available online at http://hahana.soest.hawaii.edu/hot/hot-dogs/interface.html. We also improved the testbed for the Arabian Sea and Equatorial Pacific by assimilating more data types including DOM, bacterial biomass and production, which were collected online at http://usjgofs.whoi.edu/jg/dir/jgofs/.

3.1. Forward Model Setup

Starting from the initial condition of the state variables, the model computes the temporal evolution of the state variables until the end of modeling period using a constant time step of 1 hour. The temporal evolution is affected by (1) ecological derivatives defined by the equations described in Section 2, which also requires inputs of PAR and temperature from physical forcings; and (2) physical derivatives from vertical advection, vertical diffusion, bottom boundary exchange and sinking of detritus, which requires inputs of vertical velocity, vertical diffusive coefficient, mixed layer depth and sinking speed of detritus. Vertical advection and detrital sinking are described with a third-order direct space-time upwind-biased scheme (Hundsdorfer & Trompert 1994) and the Sweby flux limiter (Sweby 1984) and are simplified to work for 1-D (vertical) advection only. Vertical diffusion is applied using a Crank-Nicholson vertically variable diffusion operation (Press et al. 1986), with a closed upper boundary and an open bottom boundary. The effect of the horizontal advection of biological quantities (only applying to the Equatorial Pacific site as discussed below) is examined through a scaling analysis using output from a 3-D coupled biological–physical models (Christian et al. 2002). The ecological derivatives are not related to space differentials, and a simple forward-time scheme is used to solve the ecological model.

3.1.1. Model Grid and Stations

The testbed is a 1-D vertical frame that allows users to define a vertical grid per their own applications. Our application uses 25 layers, with 10 surface layers of 5 m each and 15 bottom layers of 10 m each. So the model domain covers the upper 200 m water

column. The testbed allows users to define the length of time steps. Our application uses 1 hour for each time step because it is short enough to resolve diurnal patterns and long enough that the outputs never showed any instability.

The main study site is the Hawaii Ocean Time-series study station ALOHA (22° 45′ N, 158° W, HOT hereafter) is 100 km north of Oahu, Hawaii with bottom depth of 4700 m. There are two setups for this site. The first setup continuously simulates in 1991 – 2005 with one setup of initial conditions and sequential boundary conditions and physical forcings. This setup is useful to study the accumulating affects: how the dynamics of one year can affect the next year(s). The second setup treats each year in 1990 – 2005 as a single entity with its own initial and boundary conditions and physical forcings. This setup provides more flexibility; parameters can be optimized either for a specific year or for several years simultaneously.

We will also use two supplementary sites from the Arabian Sea (AS) in 1994 and Equatorial Pacific (EQP) in 1992 to compare the differences among them and Station ALOHA. The testbed for these two site have already been setup by others (Friedrichs et al. 2007), and we add more assimilative observation types to the testbed. It only covers upper 150 m at these two sites with 10 surface layers of 5 m each and 10 bottom layers of 10 m each.

3.1.2. Initial and boundary Conditions

Initial conditions are set up at 90 days before the beginning of each calendar year (October 3rd of previous year) so that a 90-day spin-up is conducted before the simulation to reduce the influence of initial conditions on the outputs. For the continuous setup, the initial conditions are setup at October 3rd, 1989 so that 455-day spin-up is

conducted. The initial conditions are estimated by linearly interpolating in time two HOT profiles immediately before and after that date if possible.

Here we discuss how the initial conditions are estimated from observations. Please note same method are used to estimate assimilative data if applicable. Phytoplankton biomass is estimated from numbers of *Prochlorococcus*, *Synechococcus* and nano-eukaryotes, with total N biomass estimated with conversion factor values of 9.5 femtogram N $cell^{-1}$ for *Prochlorococcus*, 50 femtogram N $cell^{-1}$ for *Synechococcus* (Bertilsson et al. 2003), and 400 femtogram N $cell^{-1}$ for nano-eukaryotes. The study of Bertilsson et al. (2003) showed P limitation can change cellular C and P composition but has no effect on N composition, which is the reason we estimated N biomass instead of C or P from the cell counts and only assimilate N biomass when conducting data assimilation. There is no clear conversion factor for eukaryotes because of their diversity. The number used here assumes an average diameter ~3.5 times (~40 times in volume) that of *Prochlorococcus* and a N concentration comparable to *Prochlorococcus*. Then Redfield ratio is used to derive C and P biomass from the estimated N biomass. The initial phytoplankton biomass was assumed to consist of 80% PHY, 10% TR and 10% UN. As there are no measurements of protozoan zooplankton, PRT biomass was estimated to be 60% of total phytoplankton biomass including PHY, TR and UN. The C biomass of BA is estimated from HOT heterotrophic bacteria counts using carbon conversion factor of 10 femtogram C $cell^{-1}$ (Christian & Karl 1994, Ducklow 2000). A molar C:N:P ratio of 50:9:1 (Goldman et al. 1987, Fagerbakke et al. 1996) was used to derive N and P biomass of BA from estimated C biomass. Zooplankton tows (200 μm mesh) in the surface ocean (variable depth, on average upper 150 m) provide the

integrated C and N MZ biomass in 5 size segments. Most cruises provided two data sets, one during day and one during night. Because our model does not represent diurnal vertical migration of zooplankton, we used the averaged value from the day and night tows for each cruise. The model's intrinsic requirement of balanced relationship between MZ and phytoplankton may be different from observations, and there are uncertainties related to the estimated phytoplankton and MZ biomass as discussed above. Therefore the testbed uses the average of the zooplankton data over all the cruises to construct initial C and N biomass of MZ by assuming they are uniformly distributed vertically, and the initial P biomass of MZ by assuming a molar N:P of 25:1 (Elser & Hassett 1994). Therefore every year has the same initial conditions for MZ. This setup initializes the model with reasonable MZ biomass and expects the spin-up to solve the vertical distribution of MZ as well as balance between MZ and phytoplankton. Total particulate organic C, N and P are provided by observation, but the corresponding state variables of the model include the aggregate of PHY, TR, UN, BA, PRT and DET. To construct initial conditions for DET, we assume 40% of total particulate organic matter is detritus, because the total initial concentration of PHY, TR, UN, BA and PRT on average is ~ 60% of the corresponding POM observation. SDOM was initialized by deducting refractory DOM from the total DOM stocks provided by observation. The refractory DOM level is estimated by assuming all the DOM below 2,000 m is refractory (Carlson 2002) and taking the temporal average values because theoretically the turnover time of refractory DOM is long enough to maintain a constant level. However, the observed DOC concentration below 2000 m does show significant difference before and after 2000, averaged 44.4 μM for 1993 – 2000 and 37.8 μM for 2001-2005. DOC measurements

were probably overestimated before 2000 (M.J. Church, University of Hawaii, personal communication). Thus we assumed the DOC measurements before 2000 were overestimated and they were reduced by ~ 17% (44.4μM / 37.8μM). A value of 38 μM was used to represent refractory DOC level and the semi-labile DOC was computed by difference. Refractory DON and DOP are 2.2μM and 0.029 μM respectively. When semi-labile dissolved organic C, N and P data are not available at HOT, observed average C:N:P of 221:21:1 was used to reconstruct missing data. There are no labile DOM data available. We initialized LDOM with 2% of semi-labile DOM. NO3 and PO4 were initialized directly from observations. No ammonium data are available, and 0.1μM was used for NH4 initial conditions. In summary, although the initial conditions were constructed from our best knowledge, not all the state variables are strictly initialized from observations.

As the bottom boundary is set up at 200 m depth which is far below the euphotic zone, it is safe to assume vertical gradients of most biological species are zero and the concentrations at the lower boundary are the same as the layer immediately above. Exceptions are nitrate, phosphate and semi-labile DOM. The first two substrates are major nutrients whose upward transport is crucial to fuel new primary production in the surface layer. Semi-labile DOM dominates the total organic matter pool and shows strong gradients from surface to deep water. The concentration of these substrates at the bottom boundary was set by linearly interpolating the observed data. The estimate for semi-labile DOM was same as that for initial conditions at that depth. Another nutrient, ammonium, is expected to have a gradient at that depth because organic matter is remineralized from surface to deep. However, there are no data available so we assume the ammonium level

is much lower than nitrate at the boundary, and its gradient is negligible compared to that of nitrate.

3.1.3. Physical Forcings

Sea surface PAR, mixed layer depth (MLD), water temperature (T), vertical velocity (w) and vertical eddy diffusivity (Kz) were set to force the model. PAR was specified for each time step, and the others had temporal resolution of 1 day. T, w and Kz were set up at every vertical grid point. The physical forcings for AS and EQP were set up by others (Friedrichs et al. 2007). Here we will only discuss the setup of physical forcings for HOT.

(a) PAR: is required for photosynthesis in the model. Hourly clear-sky short wave radiation at the sea surface for the station location was computed using a standard method described by Rosati & Miyakoda (1988). It was then compared to NCEP reanalysis daily average short wave radiation interpolated to the station location to estimate the effect of daily average cloud coverage and albedo in every day (NCEP reanalysis provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.cdc.noaa.gov/). Thus, the hourly shortwave radiation reaching the sea surface was estimated while the daily average levels are same as NCEP data. Then PAR was estimated by assuming that PAR is 46% of total shortwave radiation (Pinker & Laszlo 1992, Kirk 1994). The HOT project has measured noon PAR in unit of μEinstein/cm^2/s at Station ALOHA during its monthly cruises since 1998. By using 550 nm to represent the whole PAR spectrum and therefore converting the measured PAR to Watt with 1 μEinstein/s = 0.2174 Watt, the measured noon PAR at HOT is $43 \pm 12\%$ of the corresponding estimated shortwave, which is quite comparable with the 46% used here. The attenuation of PAR with depth in the water column is controlled by:

$$PAR(z) = PAR0 \cdot e^{-(kw+kchl \cdot Chl) \cdot z}, \quad (15)$$

where z is depth in m, $PAR0$ is PAR level at sea surface, $kw = 0.038$ m^{-1} is the attenuation coefficient for seawater, $kchl = 0.05$ (mg Chl)$^{-1}$ m^2 is the attenuation coefficient for chlorophyll (Chl) in mg Chl m^{-3}.

(b) MLD: The criteria we used for MLD is the depth for a temperature change of 0.5 ^{o}C from the sea surface (Levitus 1982). Another criteria also defined by Levitus (1982) is a potential density change of 0.125 kg m^{-3} from surface. But most time the difference in MLD defined by these two criteria is small at Station ALOHA (less than 10 m in 75% of cruises). We also found when large disagreements were obtained, usually the results from the first criteria are preferable. So HOT temperature data were used to compute MLD and linearly interpolated to each model time step.

(c) T: is required by temperature effects in the model. Temperature was directly obtained from observations and linearly interpolated.

(d) w: is required by vertical advection of the physical scheme. Generally w was computed by tracking the observed vertical isopycnal displacement. The first derivative of the displacement was the estimate for w at that isopycnal depth. By linearly reducing w to zero at the surface, vertical profiles of w were constructed. With observed CTD potential density data, we located a potential density anomaly (25.163 kg m^{-3}) whose corresponding depth time-series, when averaged over time, equals 200 m. Thus by tracking the location of this potential density, w was constructed at mid point between two adjacent cruises. w was linearly interpolated to each time step. However, the w obtained from observations only has monthly resolution and may not catch submesoscale features and mesoscale eddies that can be an important source of upwelled nutrients

(Benitez-Nelson et al. 2007). Siegel et al. (1999) proposed the use of satellite-derived sea level anomaly (SLA) to address the mesoscale eddies. The basic idea is that upward isopycnal displacement occurs with negative SLA for cyclonic eddies, and downward displacement relates to positive SLA for anti-cyclonic eddies. AVISO (Archiving, Validation and Interpretation of Satellite Oceanographic data) provides satellite-derived SLA from Oct 1992 to present. The altimeter products were produced by Ssalto/Duacs and distributed by AVISO, with support from CNES. The data are available online at http://www.jason.oceanobs.com/. AVISO "delayed time updated series SLA" data with 7-day intervals were used by the testbed to cover 1993 – 2005. Then linear regression was used to estimate the relationship between HOT monthly isopycnal displacements and corresponding AVISO SLA. A 2.4 m displacement per 1 cm SLA was obtained (Siegel et al. 1999 estimated ~ 4 m displacement per 1 cm SLA in Sargasso Sea). This scale factor was then applied to all SLA data points to construct a w time series with 7-day intervals, which was then linearly interpolated to each time step. However, AVISO states that the time resolution of the SLA product, while reported at 7-day intervals, only be about 2 months, which is too long to diagnose mesoscale eddies. AVISO also provides "real time SLA", which provides data with 7-day, 3-day or 1-day intervals at different periods over 2002-2005. The data source has daily resolution and may help to solve mesoscale eddies. Overall, we compute w estimated from observed CTD data in 1990-2005; w estimated from low-resolution AVISO SLA in 1993 to 2005 and w estimated from high-resolution AVISO SLA for 2002 to 2005. Our analysis showed that low-resolution SLA can increase total positive upwelling by ~5% and high-resolution SLA can increase positive upwelling by 43% compared to CTD observations.

The w estimated here may leads to neglect the transport incurred by Ekman pumping (Doney et al. 1996, Williams & Follows 1998). But it may not dominate the total nutrient transport (Siegel et al. 1999), and our model does not solve it.

(e) Kz: is required for vertical diffusion in the physical scheme. The physical scheme assumes that all the state variables in the mixed layer are instantaneously mixed by averaging all the state variables in the mixed layer at every time step. Thus Kz values above MLD are not required. Below the MLD, vertical eddy diffusivity was constructed by reducing Kz vertically:

$$Kz(z) = Kz0 \cdot e^{-\alpha(z-MLD)}, \tag{16}$$

where z is depth in m (positive) below MLD, $Kz0 = 1.1\times10^{-4}$ m^2s^{-1} is the Kz at the bottom of mixed layer and $\alpha = 0.01$. This interpolation was suggested by J.R. Christian (University of Victoria, personal communication). The average Kz at 200 m obtained from equation (15) is ~ 3.0×10^{-5} m^2s^{-1}, which is within the range of the estimates at Station ALOHA by Christian et al. (1997).

3.2. Assimilative Data and Cost Function

We included as many data types as possible in the assimilative data list whenever an observed data type could be related to a corresponding modeling output. Table 3 shows the 17 data types used for assimilation and their availability for each year. Time series of each data type was constructed from the available observations, according to the notes below.

(a) Nutrients: nitrate and phosphate. Because of the steep gradient of inorganic

nutrients across the mixed layer depth and low or even effectively no measurable inorganic nutrients in surface waters, it is a challenge for the model to fit the nutrient profiles. The nutrient level below mixed layer is largely determined by physical instead of ecological processes. In addition, the mixed layer depth is usually comparable or even deeper than the depth of euphotic zone. Thus the nutrient level below the mixed layer depth has less influence on the primary production compared to the nutrient level at surface. Even if primary production is significant below the mixed layer, the high nutrient level allows a larger error considering it is high enough to remove limitation on primary production. Therefore, modeling the correct nutrient level in mixed layer is relatively more important. Thus only the average values of nutrients in mixed layer of each profile were assimilated and one data point was put at 10 m depth to make sure it is inside the mixed layer. Again, the physical scheme of the model mixes all materials in the mixed layer and this single point is expected to constrain the nutrient level of the whole mixed layer. At HOT site where the nitrate level is low in the mixed layer, we used the high-sensitivity low-level nitrate data for this study.

(b) Phytoplankton N biomass: was estimated the same way as for the initial conditions, and the model equivalent is total N biomass of PHY and UN. TR is not included as the phytoplankton biomass was estimated from cell counts, and the colonial *Trichodesmium* spp. were not counted routinely.

(c) Phytoplankton Chlorophyll *a*: was obtained directly from the fluorometric data at HOT, and from fluorometric or HPLC (high performance liquid chromatography) data at AS and EQP depending on availability. The model equivalent is the total Chlorophyll *a* of PHY, TR and UN.

(d) Primary production: its model equivalent is gross C growth of PHY, UN and TR minus their DOM excretion because the ^{14}C tracer method used for the primary production measurements largely captures only the particulate production through filtration. But it is possible that ^{14}C-labeled dissolved organic carbon (DOC) can adsorb onto glass fiber filters, so that the measured PP may also include portion of dissolved PP (Karl et al. 1998). As the DOC that included in the measurement is difficult to be quantified, we still only use particulate PP as the model equivalent, which may potentially lead to overestimation of PP.

(e) Heterotrophic bacterial C biomass and production: biomass was estimated the same way as for the initial conditions. Heterotrophic bacterial C production rates at HOT were provided by M.J. Church at University of Hawaii including 14 vertical profiles of leucine incorporation rates in 2000, 2001, 2002, 2004 and 2005. A conversion factor of 1.5 kg C /mol leucine (Ducklow 2000) was used to derive production rates. Heterotrophic bacterial production at HOT is the only data type used here that is not publicly available (yet). Heterotrophic bacterial C production rates at AS and EQP were calculated using same conversion factor from leucine incorporation rates.

(f) Mesozooplankton C biomass: was estimated the same way as initial conditions. Again, the averaged profile of the day and night tows were used for each cruise. But because there was only one depth-integrated tow for each profile, the volume averaged value of each profile was put at 10 m depth in order to ensure it is inside the mixed layer. For 1990 – 1993 when the data were not measured, seasonal average biomass data for 1994 – 1997 were used. Observations in 1999 – 2001 were used the same way to construct data for 2002 -2005. The model equivalent of mesozooplankton biomass is

metazoan C biomass.

(g) Semi-labile DOC, DON and DOP: were estimated the same way as initial conditions. DOC data were not measured in 1990 - 1992 and DON and DOP were not measured in 2002 - 2005. The missing data were derived from the other available components of DOM using the observed average DOM C:N:P of 220:19:1 as described for initial conditions.

(h) Particulate organic C, N and P: their model equivalents are total biomass of PHY, TR, UN, PRT and DET according to the POM collecting method used by HOT program. MZ and BA are not included because large zooplankton were removed by 202 μm mesh before the collection and the combusted GF/F glass fiber filters (0.7 μm nominal pore width) may not be able to collect heterotrophic bacteria.

(i) Sediment trap C, N and P data: their model equivalents are DET sinking flux out of the layer immediately above the sediment trap. Physical transportation of DET by advection or diffusion is not included.

The cost function J is a metric to evaluate the difference between modeled results a and observations $\hat{a}$. Lower J values indicate better fits of the model to observations. The cost function J used here summarizes the average Chi-squared value of each data type:

$$J = \sum_{m=1}^{M} \frac{1}{N_{s,m}} \sum_{n=1}^{N_m} \left(\frac{a_{m,n} - \hat{a}_{m,n}}{\sigma_m} \right)^2 , \tag{17}$$

where m and n represent assimilative data type and data points, M and N are the total numbers of assimilative data types and data points, σ is the target error which is defined as

$$\sigma = \overline{\hat{a}_n} \cdot \frac{e}{q}, \tag{18}$$

where q is the adjustment factor which will be discussed later, $\overline{\hat{a}_n}$ is the mean of observations, and e is fractional error level of the observation. To reasonably estimate e, we assumed all materials in the mixed layer should be completely mixed and a perfect measurement without any error should give a profile with same data value at every measured depth in the mixed layer. Thus we can use the ratio of standard deviation to mean value (coefficient of variation CV) of each vertical profile in the mixed layer to represent the error level for that measurement. The averaged CV of all the profiles was used as e and each data type has a different e value for each site. This method does not apply to mesozooplankton and sediment trap data because there is only one data point per profile for them. Their fractional error level was assumed to be 25%.

The adjustment factor q was used to change the target error with adjusting the fractional error level to e/q (Chapter 3, Table 1) when considering the issues discussed below. Additional uncertainties are related to conversion factors for estimating phytoplankton biomass, bacterial biomass and production so that the target errors for these variables were increased accordingly. Primary production is not mixed in the mixed layer and therefore its error level was reduced. But the CV of the primary production at EQP was already very low (6.7%) thus no adjustment was made to it. The adjustment factors for C/N/P components of DOM, POM and sinking particles (sediment traps) in HOT are set to $(1/3)^{1/2}$ so that the total contribution of these 3 C/N/P components is considered to be one data type. Similar treatment was made to AS and EQP. In addition, we observed very low CVs (< 5%) for nitrate, phosphate and DOC in EQP so that they

were increased in order to avoid fitting the model to noise.

3.3. Adjoint Model

We used auto-differential software to conduct code-to-code translation to automatically generate adjoint codes for the forward model. The ecosystem model was translated by TAPENADE 2.1 (http://www-sop.inria.fr/tropics/tapenade.html) developed by Institut National de Recherche en Informatique en Automatique, France (INRIA). However, the original regional testbed framework, including the physical scheme and integration, used TAMC developed by Giering & Kaminski (http://www.autodiff.com/tamc/) to generate adjoint code. We used TAPENADE instead of TAMC because TAPENADE can be run locally while TAMC requires the remote access and may become unavailable occasionally. It is also found that TAPENADE is sometimes more reliable than TAMC for our work. The performance of the two programs did not show much difference.

Nevertheless, both programs are based on same scheme, which we briefly describe in Appdendix B. The description is based on Lawson et al. (1995), Hascoët & Pascual (2004) and Giering (1999).

3.4. Optimization Procedure: Limited-memory Quasi-Newton Method

The testbed includes a Fortran program, M1QN3 3.1, developed by J.C. Gilbert & C. Lemaréchal, which uses the gradient information calculated from the adjoint code to minimize the cost function. It is based on a limited-memory quasi-Newton scheme first proposed by Nocedal (1980) and then improved by Gilbert & Lemaréchal (1989). The

method is briefly described in Appendix C based on the manual of M1QN3 and a note prepared by Galen Andrew for an informal tutorial at Microsoft Research (http://research.microsoft.com/users/galena/). This scheme is unconstrained and there is no penalty term to the cost function. Thus the parameters are basically optimized freely. But the parameters were logarithmically transformed before they were submitted to optimization, and transformed back exponentially after the optimization. By this way all the parameters are positive after the optimization.

The potential limitation of this optimization method is that theoretically it can only reach local minima, although we seek the solution for the global minimum. For an ecosystem application, most parameters can be initialized in a reasonable range based on literature search and rationality of ecological meanings, which may help to avoid local minima.

3.5. Selection of Optimizable Parameters and Uncertainty Analysis

Mathematically when the errors in the observations are assumed to be normally distributed, the final inverse Hessian matrix of the least-square cost function at its minimum can be used to analyze the uncertainties (one standard deviation) of the final optimized parameter values (Tziperman & Thacker 1989, Matear 1995), which approximately applies to our cost function. Denoting x as the realized optimal solution and x^* as the real solution then $\left(x - x^*\right)$ is the error of the parameters. Expanding the cost function $J\left(x\right)$ about x^* yields

$$J \approx J_{\min} + 1/2\left(x - x^*\right)^T B\left(x - x^*\right), \tag{19}$$

which to the first order equals to zero because the cost function has zero gradient with respect to all parameters at the realized optimal solution, and the third and higher-order terms are neglected. If the neglected terms are sufficiently small, then the uncertainties $\left(x - x^*\right)$ are normally distributed with zero mean and with an error-covariance matrix defined as the inverse of the Hessian, $H = B^{-1}$ (Tziperman & Thacker 1989). The error-covariance matrix provides information on the probability distribution of the optimal parameters. The square roots of diagonal elements of the matrix provide a measure of the width of the distribution for the different optimal parameters and represent the uncertainties of parameters. Therefore an optimized parameter is well constrained when its corresponding diagonal element in the inverse Hessian has low value. Also other elements not on the diagonal represent the covariance of the corresponding two parameters. A high value indicates the corresponding two parameters are highly correlated and can not be simultaneously optimized because a change to one of the parameters will be counteracted by a change to the other with multiple pairs of parameter values producing indistinguishable results. As the Hessian matrix typically cannot be solved explicitly, a finite difference approximation is commonly used to generate the Hessian matrix.

In summary, an approximate inverse Hessian matrix H has to be computed after the optimization, and parameters are judged to be optimizable only when the values of their corresponding diagonal elements in H are low and no pair of parameters selected has high covariance, as indicated by their corresponding non-diagonal elements in H. The uncertainty of an optimized parameter is the square root of its diagonal element in H.

3.6. Default Parameter Values and Optimization

The default parameter values derived from a preliminary optimization are shown in Table 2. Initial values for this optimization were drawn from the literature and previous experience with the model. Two steps were conducted during the optimization.

In the first step, the gradient of the cost function with respect to each parameter was computed using the initial guesses for parameter values with all the years included simultaneously. Those parameters with small gradients were excluded from optimization as the cost function was not sensitive to them. As discussed above, a cost value of more than 1 at one single year indicates a significant difference. Here we included all the 16 years, thus a cost value of 16 indicates a significant difference. Therefore we assume if a 100% change of a parameter cannot alter the total cost by 16 according to its gradient, this parameter is considered insensitive. The remaining parameters were submitted to optimization over all the years. The output optimized parameters were subject to the inverse Hessian matrix test as described in Section 3.5, and those with highest uncertainties were excluded. Optimization was conducted iteratively with some parameters excluded at each step until all the remaining parameters passed the inverse Hessian matrix test. By this method, only a small portion of the parameters, about 10 from all 80 parameters, was selected. These parameters were well constrained and formed our “optimizable list”.

However, although other parameters were determined to be not optimizable because the assimilative data were not sufficient to constrain them, their change could also reduce the cost function. So the second step was carried out to recover this discarded information.

By optimizing all the parameters against model integration of all the years simultaneously, some of the parameters may end up with scientifically unreasonable values. These parameters were kept at their original values, and the remaining parameters were submitted to optimization again. This process was iterated until no more parameters were optimized to be scientifically unreasonable. Then default values were replaced with the new optimized values for all the parameters except those in the “optimizable list” determined by the first step, as the future studies will optimize them again anyway.

Appendix A: Ecosystem Model Equations

Notes:

(a) All the parameter names (Table 2) begin with lower-case letters, and state and local variables begin with upper-case letters.

(b) State variables are C, N and P representing carbon, nitrogen and phosphorus components respectively of a biological species or a chemical substrate as noted in the subscript.

(c) Q_X^Y is used to represent a ratio X/Y.

1. Temperature effect

(A1.1) $$Tfun = EXP\left(-A_E\left(\frac{1}{T+273.15}-\frac{1}{25+273.15}\right)\right),$$

where T is water temperature in Celsius.

2. Phytoplankton (PHY) Processes

(A2.1) Cellular quota limitation on primary production

Cellular quota

$$Q^C_{N,PHY} = \frac{N_{PHY}}{C_{PHY}}, \quad Q^C_{P,PHY} = \frac{P_{PHY}}{C_{PHY}}$$

Limitation

$$Nfun_{PHY} = \frac{Q^C_{N,PHY} - q^C_{N,\min,PHY}}{q^C_{N,rdf,PHY} - q^C_{N,\min,PHY}}, \text{ then set } \ 0 \le Nfun_{PHY} \le 1$$

$$Pfun_{PHY} = \frac{Q^C_{P,PHY} - q^C_{P,\min,PHY}}{q^C_{P,rdf,PHY} - q^C_{P,\min,PHY}}, \text{ then set } \ 0 \le Pfun_{PHY} \le 1$$

Maximum primary production rate after cell quota limitation

$$P^C_{\max,PHY} = \mu_{PHY} \cdot Tfun \cdot \min\left(Nfun_{PHY}, Pfun_{PHY}\right)$$

(A2.2) Light limitation and carbon specific gross primary production

$GROW^C_{PHY} = 0$, if $P^C_{\max,PHY} = 0$. Otherwise:

$$GROW^C_{PHY} = C_{PHY} \cdot P^C_{\max,PHY} \cdot \left(1 - EXP\left(\frac{-\alpha^{CHL}_{PHY} \cdot Q^C_{CHL,PHY} \cdot PAR}{P^C_{\max,PHY}}\right)\right) \cdot EXP\left(-\beta_{PHY} \cdot PAR\right)$$

(A2.3) Nitrogen assimilation, including nitrate uptake and ammonium uptake

Cellular quota limitation

$$V^N_{\max,PHY} = \frac{q^C_{N,\max,PHY} - Q^C_{N,PHY}}{q^C_{N,\max,PHY} - q^C_{N,rdf,PHY}}, \text{ then set } \ 0 \le V^N_{\max,PHY} \le 1$$

Nitrogen assimilation

$$GROW_{PHY}^{NH4} = C_{PHY} \cdot v_{REF,PHY}^{N} \cdot Tfun \cdot V_{\max,PHY}^{N} \cdot \frac{NH4}{NH4 + k_{PHY}^{NH4} + NO3 \cdot k_{PHY}^{NH4} / k_{PHY}^{NO3}}$$

$$GROW_{PHY}^{NO3} = C_{PHY} \cdot v_{REF,PHY}^{N} \cdot Tfun \cdot V_{\max,PHY}^{N} \cdot \frac{NO3}{NO3 + k_{PHY}^{NO3} + NH4 \cdot k_{PHY}^{NO3} / k_{PHY}^{NH4}}$$

$$GROW_{PHY}^{N} = GROW_{PHY}^{NH4} + GROW_{PHY}^{NO3}$$

(A2.4) Phosphorous (phosphate) assimilation

$$V_{\max,PHY}^{P} = \frac{q_{P,\max,PHY}^{C} - Q_{P,PHY}^{C}}{q_{P,\max,PHY}^{C} - q_{P,rdf,PHY}^{C}}, \text{ then set } \; 0 \le V_{\max,PHY}^{P} \le 1$$

$$GROW_{PHY}^{P} = C_{PHY} \cdot v_{REF,PHY}^{P} \cdot Tfun \cdot V_{\max,PHY}^{P} \cdot \frac{PO4}{PO4 + k_{PHY}^{PO4}}$$

(A2.5) Respiration required by nitrate reduction

$$RESP_{PHY} = \zeta^{NO3} \cdot GROW_{PHY}^{NO3}$$

(A2.6) Chlorophyll production

$GROW_{PHY}^{CHL} = 0$, if $PAR \le 0$. Otherwise:

$$GROW_{PHY}^{CHL} = \theta \cdot GROW_{PHY}^{N} \cdot \frac{GROW_{PHY}^{C}}{\alpha_{PHY}^{CHL} \cdot CHL_{PHY} \cdot PAR \cdot EXP\left(-\beta_{PHY} \cdot PAR\right)}$$

(A2.7) DOM Excretion

(a) Passive excretion

$$EXCR_{PSV,PHY}^{C} = ex_{PSV,PHY} \cdot C_{PHY}$$

$$EXCR_{PSV,PHY}^{N} = ex_{PSV,PHY} \cdot N_{PHY}$$

$$EXCR_{PSV,PHY}^{P} = ex_{PSV,PHY} \cdot P_{PHY}$$

(b) Active carbohydrate excretion

$$EXCR_{CHO,PHY}^{C} = ex_{CHO,PHY} \cdot GROW_{PHY}^{C}$$

(c) Active DOM release in order to adjust cellular stoichiometry

$$EXCR_{ACT,PHY}^{C} = 0.5 \cdot C_{PHY} \cdot \max\left(0,\ 1 - \frac{Q_{N,PHY}^{C}}{q_{N,rdf,PHY}^{C}},\ 1 - \frac{Q_{P,PHY}^{C}}{q_{P,rdf,PHY}^{C}}\right)$$

if $EXCR_{ACT,PHY}^{C} > 0$

$$EXCR_{ACT,PHY}^{N} = 0.5 \cdot 0.25 \cdot N_{PHY} \cdot \max\left(0,\ 1 - \frac{Q_{P,PHY}^{N}}{q_{P,rdf,PHY}^{C} / q_{N,rdf,PHY}^{C}}\right)$$

$$EXCR_{ACT,PHY}^{P} = 0.5 \cdot 0.25 \cdot P_{PHY} \cdot \max\left(0,\ 1 - \frac{Q_{N,PHY}^{P}}{q_{N,rdf,PHY}^{C} / q_{P,rdf,PHY}^{C}}\right)$$

else

$$EXCR_{ACT,PHY}^{N} = 0, \quad EXCR_{ACT,PHY}^{P} = 0$$

end

(d) Partitioning between labile and semi-labile DOM

$$EXCR_{PHY}^{LDOC} = EXCR_{PSV,PHY}^{C} + 0.75 \cdot EXCR_{CHO,PHY}^{C}$$
$$EXCR_{PHY}^{LDON} = EXCR_{PSV,PHY}^{N}$$
$$EXCR_{PHY}^{LDOP} = EXCR_{PSV,PHY}^{P}$$
$$EXCR_{PHY}^{SDOC} = EXCR_{ACT,PHY}^{C} + 0.25 \cdot EXCR_{CHO,PHY}^{C}$$
$$EXCR_{PHY}^{SDON} = EXCR_{ACT,PHY}^{N}$$
$$EXCR_{PHY}^{SDOP} = EXCR_{ACT,PHY}^{P}$$

(A2.8) POM production by aggregation

$$POM_{PHY}^{C} = pom_{PHY} \cdot {C_{PHY}}^{2}$$
$$POM_{PHY}^{N} = Q_{N,PHY}^{C} \cdot POM_{PHY}^{C}$$
$$POM_{PHY}^{P} = Q_{P,PHY}^{C} \cdot POM_{PHY}^{C}$$
$$POM_{PHY}^{CHL} = Q_{CHL,PHY}^{C} \cdot POM_{PHY}^{C}$$

(A2.9) Grazing on phytoplankton by protozoa

$$GRAZ_{PHY}^{C} = Tfun \cdot \mu_{PRT} \cdot C_{PRT} \cdot \frac{{C_{PHY}}^{2}}{{C_{PHY}}^{2} + {g_{PHY}}^{2} + \left(C_{UN} \cdot g_{PHY} / g_{UN}\right)^{2} + \left(C_{BA} \cdot g_{PHY} / g_{BA}\right)^{2}}$$
$$GRAZ_{PHY}^{N} = Q_{N,PHY}^{C} \cdot GRAZ_{PHY}^{C}$$
$$GRAZ_{PHY}^{P} = Q_{P,PHY}^{C} \cdot GRAZ_{PHY}^{C}$$
$$GRAZ_{PHY}^{CHL} = Q_{CHL,PHY}^{C} \cdot GRAZ_{PHY}^{C}$$

(A2.10) PHY net growth rate

$$\frac{dC_{PHY}}{dt} = GROW_{PHY}^{C} - RESP_{PHY} - EXCR_{PHY}^{LDOC} - EXCR_{PHY}^{SDOC} - POM_{PHY}^{C} - GRAZ_{PHY}^{C}$$
$$\frac{dN_{PHY}}{dt} = GROW_{PHY}^{N} - EXCR_{PHY}^{LDON} - EXCR_{PHY}^{SDON} - POM_{PHY}^{N} - GRAZ_{PHY}^{N}$$
$$\frac{dP_{PHY}}{dt} = GROW_{PHY}^{P} - EXCR_{PHY}^{LDOP} - EXCR_{PHY}^{SDOP} - POM_{PHY}^{P} - GRAZ_{PHY}^{P}$$
$$\frac{dCHL_{PHY}}{dt} = GROW_{PHY}^{CHL} - POM_{PHY}^{CHL} - GRAZ_{PHY}^{CHL}$$

3. *Trichodesmium* spp. (TR) Processes

Notes: Trichodesmium is disabled for the simulations at AS and EQP

(A3.1) Cellular quota limitation on primary production

Limitation

$$Nfun_{TR} = \frac{Q^C_{N,TR} - q^C_{N,\min,TR}}{q^C_{N,rdf,TR} - q^C_{N,\min,TR}}, \text{ then set } \; 0 \le Nfun_{TR} \le 1$$

$$Pfun_{TR} = \frac{Q^C_{P,TR} - q^C_{P,\min,TR}}{q^C_{P,rdf,TR} - q^C_{P,\min,TR}}, \text{ then set } \; 0 \le Pfun_{TR} \le 1$$

Maximum primary production rate after cell quota limitation

$$P^C_{\max,TR} = \mu_{TR} \cdot Tfun \cdot \min\left(Nfun_{TR}, Pfun_{TR}\right)$$

(A3.2) Light limitation and carbon specific gross primary production

$GROW^C_{TR} = 0$, if $P^C_{\max,TR} = 0$. Otherwise:

$$GROW^C_{TR} = C_{TR} \cdot P^C_{\max,TR} \cdot \left(1 - EXP\left(\frac{-\alpha^{CHL}_{TR} \cdot Q^C_{CHL,TR} \cdot PAR}{P^C_{\max,TR}}\right)\right) \cdot EXP\left(-\beta_{TR} \cdot PAR\right)$$

(A3.3) Nitrogen assimilation, including nitrogen fixation, nitrate uptake and ammonium uptake

(a) Cellular quota limitation

$$V^N_{\max,TR} = \frac{q^C_{N,\max,TR} - Q^C_{N,TR}}{q^C_{N,\max,TR} - q^C_{N,rdf,TR}}, \text{ then set } \; 0 \le V^N_{\max,TR} \le 1$$

(b) Nitrate and ammonium uptake

$$GROW^{NH4}_{TR} = C_{TR} \cdot v^N_{REF,TR} \cdot Tfun \cdot V^N_{\max,TR} \cdot \frac{NH4}{NH4 + k^{NH4}_{TR} + NO3 \cdot k^{NH4}_{TR} / k^{NO3}_{TR}}$$

$$GROW^{NO3}_{TR} = C_{TR} \cdot v^N_{REF,TR} \cdot Tfun \cdot V^N_{\max,TR} \cdot \frac{NO3}{NO3 + k^{NO3}_{TR} + NH4 \cdot k^{NO3}_{TR} / k^{NH4}_{TR}}$$

(c) Maximum nitrogen fixation

$NFIX_{\max,TR} = 0$, if $T \le 15°C$. Otherwise:

$$NFIX_{\max,TR} = \frac{\left(GROW^C_{TR} - \zeta^{NO3} \cdot GROW^{NO3}_{TR}\right) \cdot q^C_{N,\max,TR} - GROW^{NO3}_{TR} - GROW^{NH4}_{TR}}{1 + \zeta^{NFIX} \cdot q^C_{N,\max,TR}} \cdot Tfun \cdot V^N_{\max,TR},$$

then set $NFIX_{\max,TR} \ge 0$

(d) Total nitrogen assimilation

$$GROW^N_{TR} = \min\left(C_{TR} \cdot v^N_{REF,TR} \cdot Tfun \cdot V^N_{\max,TR}, \;\; GROW^{NH4}_{TR} + GROW^{NO3}_{TR} + NFIX_{\max,TR}\right)$$

(e) Actual nitrogen fixation

$$GROW^{NFIX}_{TR} = GROW^N_{TR} - GROW^{NH4}_{TR} - GROW^{NO3}_{TR}$$

(A3.4) Phosphorous assimilation

(a) Phosphate assimilation

$$V^P_{\max,TR} = \frac{q^C_{P,\max,TR} - Q^C_{P,TR}}{q^C_{P,\max,TR} - q^C_{P,rdf,TR}}, \text{ then set } \; 0 \le V^P_{\max,TR} \le 1$$

$$GROW_{TR}^{PO4} = C_{TR} \cdot v_{REF,TR}^{P} \cdot Tfun \cdot V_{\max,TR}^{P} \cdot \frac{PO4}{PO4 + k_{TR}^{PO4}}$$

(b) Active phosphate taking-up from deep water

$$PICK_{TR}^{PO4} = \max\left(0,\ pick_{TR}^{PO4} \cdot \left(\left(q_{P,\max,TR}^{C} + q_{P,rdf,TR}^{C}\right)/2 \cdot C_{TR} - P_{TR}\right)\right)$$

(c) Total phosphorus growth

$$GROW_{TR}^{P} = GROW_{TR}^{PO4} + PICK_{TR}^{PO4}$$

(A3.5) Respiration required by nitrogen fixation and nitrate reduction

$$RESP_{TR} = \zeta^{NO3} \cdot GROW_{TR}^{NO3} + \zeta^{NFIX} \cdot GROW_{TR}^{NFIX}$$

(A3.6) Chlorophyll production

$GROW_{TR}^{CHL} = 0$, if $PAR \le 0$. Otherwise:

$$GROW_{TR}^{CHL} = \theta \cdot GROW_{TR}^{N} \cdot \frac{GROW_{TR}^{C}}{\alpha_{TR}^{CHL} \cdot CHL_{TR} \cdot PAR \cdot EXP\left(-\beta_{TR} \cdot PAR\right)}$$

(A3.7) DOM and ammonium excretion

(a) Passive excretion

$$EXCR_{PSV,TR}^{C} = ex_{PSV,TR} \cdot C_{TR}$$

$$EXCR_{PSV,TR}^{N} = ex_{PSV,TR} \cdot N_{TR}$$

$$EXCR_{PSV,TR}^{P} = ex_{PSV,TR} \cdot P_{TR}$$

(b) Active carbohydrate excretion

$$EXCR_{CHO,TR}^{C} = ex_{CHO,TR} \cdot GROW_{TR}^{C}$$

(c) Active DOM release in order to adjust cellular stoichiometry

$$EXCR_{ACT,TR}^{C} = 0.5 \cdot C_{TR} \cdot \max\left(0,\ 1 - \frac{Q_{N,TR}^{C}}{q_{N,rdf,TR}^{C}},\ 1 - \frac{Q_{P,TR}^{C}}{q_{P,rdf,TR}^{C}}\right)$$

if **Error! Objects cannot be created from editing field codes.**

$$EXCR_{ACT,TR}^{N} = 0.5 \cdot 0.25 \cdot N_{TR} \cdot \max\left(0,\ 1 - \frac{Q_{P,TR}^{N}}{q_{P,rdf,TR}^{C} / q_{N,rdf,TR}^{C}}\right)$$

$$EXCR_{ACT,TR}^{P} = 0.5 \cdot 0.25 \cdot P_{TR} \cdot \max\left(0,\ 1 - \frac{Q_{N,TR}^{P}}{q_{N,rdf,TR}^{C} / q_{P,rdf,TR}^{C}}\right)$$

else

$$EXCR_{ACT,TR}^{N} = 0, \quad EXCR_{ACT,TR}^{P} = 0$$

end

(d) Active DON release of newly fixed N

$$EXCR_{NFIX,TR}^{DON} = 0.5 \cdot ex_{NFIX,TR} \cdot GROW_{TR}^{NFIX} \cdot Nfun_{TR}$$

(e) Partitioning between labile and semi-labile DOM

$$EXCR_{TR}^{LDOC} = EXCR_{PSV,TR}^{C} + 0.75 \cdot EXCR_{CHO,TR}^{C}$$

$$EXCR_{TR}^{LDON} = EXCR_{PSV,TR}^{N} + EXCR_{NFIX,TR}^{DON}$$

$$EXCR_{TR}^{LDOP} = EXCR_{PSV,TR}^{P}$$
$$EXCR_{TR}^{SDOC} = EXCR_{ACT,TR}^{C} + 0.25 \cdot EXCR_{CHO,TR}^{C}$$
$$EXCR_{TR}^{SDON} = EXCR_{ACT,TR}^{N}$$
$$EXCR_{TR}^{SDOP} = EXCR_{ACT,TR}^{P}$$

(f) Ammonium excretion

$$EXCR_{TR}^{NH4} = 0.5 \cdot ex_{NFIX,TR} \cdot GROW_{TR}^{NFIX} \cdot Nfun_{TR}$$

(A3.8) POM production by aggregation

$$POM_{TR}^{C} = pom_{TR} \cdot C_{TR}^{2}$$
$$POM_{TR}^{N} = Q_{N,TR}^{C} \cdot POM_{TR}^{C}$$
$$POM_{TR}^{P} = Q_{P,TR}^{C} \cdot POM_{TR}^{C}$$
$$POM_{TR}^{CHL} = Q_{CHL,TR}^{C} \cdot POM_{TR}^{C}$$

(A3.9) Grazing on *Trichodesmium* by metazoa

$$GRAZ_{TR}^{C} = Tfun \cdot \mu_{MZ} \cdot C_{MZ} \cdot \frac{C_{TR}^{2}}{C_{TR}^{2} + g_{TR}^{2} + \left(C_{PRT} \cdot g_{TR} / g_{PRT}\right)^{2}}$$
$$GRAZ_{TR}^{N} = Q_{N,TR}^{C} \cdot GRAZ_{TR}^{C}$$
$$GRAZ_{TR}^{P} = Q_{P,TR}^{C} \cdot GRAZ_{TR}^{C}$$
$$GRAZ_{TR}^{CHL} = Q_{CHL,TR}^{C} \cdot GRAZ_{TR}^{C}$$

(A3.10) Trichodesmium net growth rate

$$\frac{dC_{TR}}{dt} = GROW_{TR}^{C} - RESP_{TR} - EXCR_{TR}^{LDOC} - EXCR_{TR}^{SDOC} - POM_{TR}^{C} - GRAZ_{TR}^{C}$$
$$\frac{dN_{TR}}{dt} = GROW_{TR}^{N} - EXCR_{TR}^{NH4} - EXCR_{TR}^{LDON} - EXCR_{TR}^{SDON} - POM_{TR}^{N} - GRAZ_{TR}^{N}$$
$$\frac{dP_{TR}}{dt} = GROW_{TR}^{P} - EXCR_{TR}^{LDOP} - EXCR_{TR}^{SDOP} - POM_{TR}^{P} - GRAZ_{TR}^{P}$$
$$\frac{dCHL_{TR}}{dt} = GROW_{TR}^{CHL} - POM_{TR}^{CHL} - GRAZ_{TR}^{CHL}$$

4. Unicellular N_2-fixers (UN) Processes

Notes: Unicellular N_2-fixers is disabled for the simulations at AS and EQP

(A4.1) Cellular quota limitation on primary production

Limitation

$$Nfun_{UN} = \frac{Q_{N,UN}^{C} - q_{N,\min,UN}^{C}}{q_{N,rdf,UN}^{C} - q_{N,\min,UN}^{C}}, \text{ then set } \; 0 \le Nfun_{UN} \le 1$$

$$Pfun_{UN} = \frac{Q^C_{P,UN} - q^C_{P,\min,UN}}{q^C_{P,rdf,UN} - q^C_{P,\min,UN}}, \text{ then set } \; 0 \le Pfun_{UN} \le 1$$

Maximum primary production rate after cell quota limitation

$$P^C_{\max,UN} = \mu_{UN} \cdot Tfun \cdot \min\left(Nfun_{UN}, Pfun_{UN}\right)$$

(A4.2) Light limitation and carbon specific gross primary production

$GROW^C_{UN} = 0$, if $P^C_{\max,UN} = 0$. Otherwise:

$$GROW^C_{UN} = C_{UN} \cdot P^C_{\max,UN} \cdot \left(1 - EXP\left(\frac{-\alpha^{CHL}_{UN} \cdot Q^C_{CHL,UN} \cdot PAR}{P^C_{\max,UN}}\right)\right) \cdot EXP\left(-\beta_{UN} \cdot PAR\right)$$

(A4.3) Nitrogen assimilation, including nitrogen fixation, nitrate uptake and ammonium uptake

(a) Cellular quota limitation

$$V^N_{\max,UN} = \frac{q^C_{N,\max,UN} - Q^C_{N,UN}}{q^C_{N,\max,UN} - q^C_{N,rdf,UN}}, \text{ then set } \; 0 \le V^N_{\max,UN} \le 1$$

(b) Nitrate and ammonium uptake

$$GROW^{NH4}_{UN} = C_{UN} \cdot v^N_{REF,UN} \cdot Tfun \cdot V^N_{\max,UN} \cdot \frac{NH4}{NH4 + k^{NH4}_{UN} + NO3 \cdot k^{NH4}_{UN} / k^{NO3}_{UN}}$$

$$GROW^{NO3}_{UN} = C_{UN} \cdot v^N_{REF,UN} \cdot Tfun \cdot V^N_{\max,UN} \cdot \frac{NO3}{NO3 + k^{NO3}_{UN} + NH4 \cdot k^{NO3}_{UN} / k^{NH4}_{UN}}$$

(c) Maximum nitrogen fixation

$NFIX_{\max,UN} = 0$, if $T \le 15°C$. Otherwise:

$$NFIX_{\max,UN} = \frac{\left(GROW^C_{UN} - \zeta^{NO3} \cdot GROW^{NO3}_{UN}\right) \cdot q^C_{N,\max,UN} - GROW^{NO3}_{UN} - GROW^{NH4}_{UN}}{1 + \zeta^{NFIX} \cdot q^C_{N,\max,UN}} \cdot Tfun \cdot V^N_{\max,UN},$$

then set $NFIX_{\max,UN} \ge 0$

(d) Total nitrogen assimilation

$$GROW^N_{UN} = \min\left(C_{UN} \cdot v^N_{REF,UN} \cdot Tfun \cdot V^N_{\max,UN}, \; GROW^{NH4}_{UN} + GROW^{NO3}_{UN} + NFIX_{\max,UN}\right)$$

(e) Actual nitrogen fixation

$$GROW^{NFIX}_{UN} = GROW^N_{UN} - GROW^{NH4}_{UN} - GROW^{NO3}_{UN}$$

(A4.4) Phosphorous assimilation

Phosphate assimilation

$$V^P_{\max,UN} = \frac{q^C_{P,\max,UN} - Q^C_{P,UN}}{q^C_{P,\max,UN} - q^C_{P,rdf,UN}}, \text{ then set } \; 0 \le V^P_{\max,UN} \le 1$$

$$GROW^P_{UN} = C_{UN} \cdot v^P_{REF,UN} \cdot Tfun \cdot V^P_{\max,UN} \cdot \frac{PO4}{PO4 + k^{PO4}_{UN}}$$

(A4.5) Respiration required by nitrogen fixation and nitrate reduction

$$RESP_{UN} = \zeta^{NO3} \cdot GROW_{UN}^{NO3} + \zeta^{NFIX} \cdot GROW_{UN}^{NFIX}$$

(A4.6) Chlorophyll production

$GROW_{UN}^{CHL} = 0$, if $PAR \le 0$. Otherwise:

$$GROW_{UN}^{CHL} = \theta \cdot GROW_{UN}^{N} \cdot \frac{GROW_{UN}^{C}}{\alpha_{UN}^{CHL} \cdot CHL_{UN} \cdot PAR \cdot EXP\left(-\beta_{UN} \cdot PAR\right)}$$

(A4.7) DOM and ammonium excretion

(a) Passive excretion

$$EXCR_{PSV,UN}^{C} = ex_{PSV,UN} \cdot C_{UN}$$

$$EXCR_{PSV,UN}^{N} = ex_{PSV,UN} \cdot N_{UN}$$

$$EXCR_{PSV,UN}^{P} = ex_{PSV,UN} \cdot P_{UN}$$

(b) Active carbohydrate excretion

$$EXCR_{CHO,UN}^{C} = ex_{CHO,UN} \cdot GROW_{UN}^{C}$$

(c) Active DOM release in order to adjust cellular stoichiometry

$$EXCR_{ACT,UN}^{C} = 0.5 \cdot C_{UN} \cdot \max\left(0,\ 1 - \frac{Q_{N,UN}^{C}}{q_{N,rdf,UN}^{C}},\ \ 1 - \frac{Q_{P,UN}^{C}}{q_{P,rdf,UN}^{C}}\right)$$

if $EXCR_{ACT,UN}^{N} > 0$

$$EXCR_{ACT,UN}^{N} = 0.5 \cdot 0.25 \cdot N_{UN} \cdot \max\left(0,\ \ 1 - \frac{Q_{P,UN}^{N}}{q_{P,rdf,UN}^{C} / q_{N,rdf,UN}^{C}}\right)$$

$$EXCR_{ACT,UN}^{P} = 0.5 \cdot 0.25 \cdot P_{UN} \cdot \max\left(0,\ \ 1 - \frac{Q_{N,UN}^{P}}{q_{N,rdf,UN}^{C} / q_{P,rdf,UN}^{C}}\right)$$

else

$$EXCR_{ACT,UN}^{N} = 0, \quad EXCR_{ACT,UN}^{P} = 0$$

end

(d) Active DON release of newly fixed N

$$EXCR_{NFIX,UN}^{DON} = 0.5 \cdot ex_{NFIX,UN} \cdot GROW_{UN}^{NFIX} \cdot Nfun_{UN}$$

(e) Partitioning between labile and semi-labile DOM

$$EXCR_{UN}^{LDOC} = EXCR_{PSV,UN}^{C} + 0.75 \cdot EXCR_{CHO,UN}^{C}$$

$$EXCR_{UN}^{LDON} = EXCR_{PSV,UN}^{N} + EXCR_{NFIX,UN}^{DON}$$

$$EXCR_{UN}^{LDOP} = EXCR_{PSV,UN}^{P}$$

$$EXCR_{UN}^{SDOC} = EXCR_{ACT,UN}^{C} + 0.25 \cdot EXCR_{CHO,UN}^{C}$$

$$EXCR_{UN}^{SDON} = EXCR_{ACT,UN}^{N}$$

$$EXCR_{UN}^{SDOP} = EXCR_{ACT,UN}^{P}$$

(f) Ammonium excretion

$$EXCR_{UN}^{NH4} = 0.5 \cdot ex_{NFIX,UN} \cdot GROW_{UN}^{NFIX} \cdot Nfun_{UN}$$

(A4.8) POM production by aggregation

$$POM_{UN}^{C} = pom_{UN} \cdot C_{UN}^{\ 2}$$
$$POM_{UN}^{N} = Q_{N,UN}^{C} \cdot POM_{UN}^{C}$$
$$POM_{UN}^{P} = Q_{P,UN}^{C} \cdot POM_{UN}^{C}$$
$$POM_{UN}^{CHL} = Q_{CHL,UN}^{C} \cdot POM_{UN}^{C}$$

(A4.9) Grazing on Unicellular N_2-fixers by protozoa

$$GRAZ_{UN}^{C} = Tfun \cdot \mu_{PRT} \cdot C_{PRT} \cdot \frac{C_{UN}^{\ 2}}{C_{UN}^{\ 2} + g_{UN}^{\ 2} + \left(C_{PHY} \cdot g_{UN} / g_{PHY}\right)^2 + \left(C_{BA} \cdot g_{UN} / g_{BA}\right)^2}$$
$$GRAZ_{UN}^{N} = Q_{N,UN}^{C} \cdot GRAZ_{UN}^{C}$$
$$GRAZ_{UN}^{P} = Q_{P,UN}^{C} \cdot GRAZ_{UN}^{C}$$
$$GRAZ_{UN}^{CHL} = Q_{CHL,UN}^{C} \cdot GRAZ_{UN}^{C}$$

(A4.10) Unicellular N_2-fixers net growth rate

$$\frac{dC_{UN}}{dt} = GROW_{UN}^{C} - RESP_{UN} - EXCR_{UN}^{LDOC} - EXCR_{UN}^{SDOC} - POM_{UN}^{C} - GRAZ_{UN}^{C}$$
$$\frac{dN_{UN}}{dt} = GROW_{UN}^{N} - EXCR_{UN}^{NH4} - EXCR_{UN}^{LDON} - EXCR_{UN}^{SDON} - POM_{UN}^{N} - GRAZ_{UN}^{N}$$
$$\frac{dP_{UN}}{dt} = GROW_{UN}^{P} - EXCR_{UN}^{LDOP} - EXCR_{UN}^{SDOP} - POM_{UN}^{P} - GRAZ_{UN}^{P}$$
$$\frac{dCHL_{UN}}{dt} = GROW_{UN}^{CHL} - POM_{UN}^{CHL} - GRAZ_{UN}^{CHL}$$

5. Heterotrophic Bacterial (BA) processes

(A5.1) Maximum available labile DOC

$$ALC = C_{LDOM}$$

(A5.2) Maximum available semi-labile DOC

$$ASC = C_{SDOM} \cdot \min\left(1,\ EXP\left(b_{LABI,N} \cdot \left(Q_{N,SDOM}^{C} / q_{N,BA}^{C} - 1\right)\right),\ EXP\left(b_{LABI,P} \cdot \left(Q_{P,SDOM}^{C} / q_{P,BA}^{C} - 1\right)\right)\right)$$

(A5.3) Cellular quota limitation

$Nfun_{BA} = Q_{N,BA}^{C} / q_{N,BA}^{C}$, then set $Nfun_{BA} \le 1$

$Pfun_{BA} = Q_{P,BA}^{C} / q_{P,BA}^{C}$, then set $Pfun_{BA} \le 1$

(A5.4) Bacterial gross carbon growth

Labile DOC

$$GROW_{BA}^{LDOC} = \mu_{BA} \cdot Tfun \cdot C_{BA} \cdot \min\left(Nfun_{BA,} Pfun_{BA}\right) \cdot \frac{ALC}{ALC + k_{DOM} + ASC}$$

Semi-labile DOC

$$GROW_{BA}^{SDOC} = \mu_{BA} \cdot Tfun \cdot C_{BA} \cdot \min\left(Nfun_{BA,} Pfun_{BA}\right) \cdot \frac{ASC}{ASC + k_{DOM} + ALC}$$

Total gross carbon growth

$$GROW_{BA}^{C} = GROW_{BA}^{LDOC} + GROW_{BA}^{SDOC}$$

(A5.5) Bacterial nitrogen uptake

$$GROW_{BA}^{LDON} = GROW_{BA}^{LDOC} \cdot Q_{N,LDOM}^{C}$$

$$GROW_{BA}^{SDON} = GROW_{BA}^{SDOC} \cdot \min\left(q_{N,BA}^{C},\ Q_{N,SDOM}^{C} + f_{SLCT,BA} / Nfun_{BA} \cdot \left(q_{N,BA}^{C} - Q_{N,SDOM}^{C}\right)\right)$$

$$GROW_{BA}^{NH4} = GROW_{BA}^{LDON} \cdot NH4 / N_{LDOM} / \min\left(1,\ Nfun_{BA}\right)$$

If $Nfun_{BA} < 1$ Then

$$GROW_{BA}^{NO3} = \min\left(\begin{array}{l} 0.1 \cdot NO3 \cdot \dfrac{1}{\min\left(1,\ Nfun_{BA}\right)} \cdot \dfrac{GROW_{BA}^{LDON} + GROW_{BA}^{SDON}}{N_{LDOM} + N_{SDOM}}, \\ \left(NO3 + NH4\right) \cdot \dfrac{GROW_{BA}^{LDON} + GROW_{BA}^{SDON}}{N_{LDOM} + N_{SDOM}} - GROW_{BA}^{NH4} \end{array}\right),$$

then set $GROW_{BA}^{NO3} \geq 0$

Else

$$GROW_{BA}^{NO3} = 0$$

End

$$GROW_{BA}^{N} = GROW_{BA}^{LDON} + GROW_{BA}^{SDON} + GROW_{BA}^{NH4} + GROW_{BA}^{NO3}$$

(A5.6) Bacterial phosphorus uptake

$$GROW_{BA}^{LDOP} = GROW_{BA}^{LDOC} \cdot Q_{P,LDOM}^{C}$$

$$GROW_{BA}^{SDOP} = GROW_{BA}^{SDOC} \cdot \min\left(q_{P,BA}^{C},\ Q_{P,SDOM}^{C} + f_{SLCT,BA} / Pfun_{BA} \cdot \left(q_{P,BA}^{C} - Q_{P,SDOM}^{C}\right)\right)$$

$$GROW_{BA}^{PO4} = GROW_{BA}^{LDOP} \cdot PO4 / P_{LDOM} / \min\left(1,\ Pfun_{BA}\right)$$

$$GROW_{BA}^{P} = GROW_{BA}^{LDOP} + GROW_{BA}^{SDOP} + GROW_{BA}^{PO4}$$

(A5.7) Bacterial Respiration

$$RESP_{BA} = \zeta^{NO3} \cdot GROW_{BA}^{NO3} + resp_{BA}^{B} \cdot Tfun \cdot C_{BA} + \left(resp_{\min,BA}^{A} + \left(resp_{\max,BA}^{A} - resp_{\min,BA}^{A}\right) \cdot EXP\left(-b_{RESP} \cdot GROW_{BA}^{C}\right)\right) \cdot GROW_{BA}^{C}$$

(A5.8) Refractory DOM release

$$REFR_{BA}^{C} = ex_{REFR,BA} \cdot C_{BA}$$

$$REFR_{BA}^{N} = q_{N,REFR}^{C} \cdot REFR_{BA}^{C}$$

$$REFR_{BA}^{P} = q_{P,REFR}^{C} \cdot REFR_{BA}^{C}$$

(A5.9) Inorganic nutrients regenerating / semi-labile DOM excretion

if $Q_{N,BA}^{C} > q_{N,BA}^{C}$ and $Q_{P,BA}^{C} > q_{P,BA}^{C}$ (Carbon in short, regenerate nutrients)

$$REMI_{BA}^{N} = remi_{BA} \cdot \left(N_{BA} - C_{BA} \cdot q_{N,BA}^{C}\right)$$

$$REMI_{BA}^{P} = remi_{BA} \cdot \left(P_{BA} - C_{BA} \cdot q_{P,BA}^{C} \right)$$

$$EXCR_{BA}^{SDOC} = EXCR_{BA}^{SDON} = EXCR_{BA}^{SDOP} = 0$$

elseif $Q_{N,BA}^{C} < q_{N,BA}^{C}$ and $Q_{N,BA}^{P} < q_{N,BA}^{C} / q_{P,BA}^{C}$ (Nitrogen in short, release SDOC & SDOP)

$$EXCR_{BA}^{SDOC} = ex_{ADJ,BA} \cdot \left(C_{BA} - N_{BA} / q_{N,BA}^{C} \right)$$

$$EXCR_{BA}^{SDON} = 0$$

$$EXCR_{BA}^{SDOP} = ex_{ADJ,BA} \cdot \left(P_{BA} - N_{BA} / q_{N,BA}^{C} \cdot q_{P,BA}^{C} \right)$$

$$REMI_{BA}^{N} = REMI_{BA}^{P} = 0$$

else (Phosphor in short, release SDOC & SDON)

$$EXCR_{BA}^{SDOC} = ex_{ADJ,BA} \cdot \left(C_{BA} - P_{BA} / q_{P,BA}^{C} \right)$$

$$EXCR_{BA}^{SDON} = ex_{ADJ,BA} \cdot \left(N_{BA} - P_{BA} / q_{P,BA}^{C} \cdot q_{N,BA}^{C} \right)$$

$$EXCR_{BA}^{SDOP} = 0$$

$$REMI_{BA}^{N} = REMI_{BA}^{P} = 0$$

end

(A5.10) Grazing on bacteria by protozoa

$$GRAZ_{BA}^{C} = Tfun \cdot \mu_{PRT} \cdot C_{PRT} \cdot \frac{{C_{BA}}^{2}}{{C_{BA}}^{2} + {g_{BA}}^{2} + \left(C_{PHY} \cdot g_{BA} / g_{PHY} \right)^{2} + \left(C_{UN} \cdot g_{BA} / g_{UN} \right)^{2}}$$

$$GRAZ_{BA}^{N} = Q_{N,BA}^{C} \cdot GRAZ_{BA}^{C}$$

$$GRAZ_{BA}^{P} = Q_{P,BA}^{C} \cdot GRAZ_{BA}^{C}$$

(A5.11) Viral mortality

$$MORT_{BA}^{C} = mort_{BA} \cdot C_{BA}$$

$$MORT_{BA}^{N} = mort_{BA} \cdot N_{BA}$$

$$MORT_{BA}^{P} = mort_{BA} \cdot P_{BA}$$

(A5.12) Net flux of inorganic nutrients through bacteria

$$FLUX_{BA}^{NH4} = REMI_{BA}^{N} - GROW_{BA}^{NH4}$$

$$FLUX_{BA}^{NO3} = -GROW_{BA}^{NO3}$$

$$FLUX_{BA}^{PO4} = REMI_{BA}^{P} - GROW_{BA}^{PO4}$$

(A5.13) Bacteria net growth

$$\frac{dC_{BA}}{dt} = GROW_{BA}^{C} - RESP_{BA} - REFR_{BA}^{C} - EXCR_{BA}^{SDOC} - GRAZ_{BA}^{C} - MORT_{BA}^{C}$$

$$\frac{dN_{BA}}{dt} = GROW_{BA}^{N} - REMI_{BA}^{N} - REFR_{BA}^{N} - EXCR_{BA}^{SDON} - GRAZ_{BA}^{N} - MORT_{BA}^{N}$$

$$\frac{dP_{BA}}{dt} = GROW_{BA}^{P} - REMI_{BA}^{P} - REFR_{BA}^{P} - EXCR_{BA}^{SDOP} - GRAZ_{BA}^{P} - MORT_{BA}^{P}$$

6. Protozoan (PRT) processes

(A6.1) Gross growth

$$GROW_{PRT}^{C} = GRAZ_{PHY}^{C} + GRAZ_{UN}^{C} + GRAZ_{BA}^{C}$$
$$GROW_{PRT}^{N} = GRAZ_{PHY}^{N} + GRAZ_{UN}^{N} + GRAZ_{BA}^{N}$$
$$GROW_{PRT}^{P} = GRAZ_{PHY}^{P} + GRAZ_{UN}^{P} + GRAZ_{BA}^{P}$$

(A6.2) DOM excretion

(a) Labile DOM excretion

$$EXCR_{PRT}^{LDOC} = f_{ex,PRT} \cdot ex_{PRT} \cdot GROW_{PRT}^{C}$$
$$EXCR_{PRT}^{LDON} = f_{ex,PRT} \cdot ex_{PRT} \cdot GROW_{PRT}^{N}$$
$$EXCR_{PRT}^{LDOP} = f_{ex,PRT} \cdot ex_{PRT} \cdot GROW_{PRT}^{P}$$

(b) Semi-labile DOM excretion

$$EXCR_{PRT}^{SDOC} = \left(1 - f_{ex,PRT}\right) \cdot ex_{PRT} \cdot GROW_{PRT}^{C}$$
$$EXCR_{PRT}^{SDON} = \left(1 - f_{ex,PRT}\right) \cdot ex_{PRT} \cdot GROW_{PRT}^{N} \cdot Q_{N,PRT}^{C} / q_{N,PRT}^{C}$$
$$EXCR_{PRT}^{SDOP} = \left(1 - f_{ex,PRT}\right) \cdot ex_{PRT} \cdot GROW_{PRT}^{P} \cdot Q_{P,PRT}^{C} / q_{P,PRT}^{C}$$

(A6.3) Respiration

$$RESP_{PRT} = resp_{PRT}^{B} \cdot Tfun \cdot C_{PRT} + resp_{PRT}^{A} \cdot GROW_{PRT}^{C}$$

(A6.4) Semi-labile DOM release in order to adjust stoichiometry

$$EXCR2_{PRT}^{SDOC} = ex_{ADJ,PRT} \cdot C_{PRT} \cdot \max\left(0,\ 1 - \frac{Q_{N,PRT}^{C}}{q_{N,PRT}^{C}},\ 1 - \frac{Q_{P,PRT}^{C}}{q_{P,PRT}^{C}}\right)$$
$$EXCR2_{PRT}^{SDON} = 0.5 \cdot EXCR2_{PRT}^{SDOC} \cdot Q_{N,PRT}^{C}$$
$$EXCR2_{PRT}^{SDOP} = 0.5 \cdot EXCR2_{PRT}^{SDOC} \cdot Q_{P,PRT}^{C}$$

(A6.5) Inorganic nutrient remineralization in order to adjust stoichiometry

$$REMI_{PRT}^{N} = remi_{PRT} \cdot \max\left(0,\ N_{PRT} - q_{N,PRT}^{C} \cdot C_{PRT},\ N_{PRT} - q_{N,PRT}^{C} / q_{P,PRT}^{C} \cdot P_{PRT}\right)$$
$$REMI_{PRT}^{P} = remi_{PRT} \cdot \max\left(0,\ P_{PRT} - q_{P,PRT}^{C} \cdot C_{PRT},\ P_{PRT} - q_{P,PRT}^{C} / q_{N,PRT}^{C} \cdot N_{PRT}\right)$$

(A6.6) POM production

$$POM_{PRT}^{C} = pom_{PRT} \cdot GROW_{PRT}^{C}$$
$$POM_{PRT}^{N} = q_{N,POM}^{C} \cdot POM_{PRT}^{C}$$
$$POM_{PRT}^{P} = q_{P,POM}^{C} \cdot POM_{PRT}^{C}$$

(A6.7) Grazing on protozoa by metazoa

$$GRAZ_{PRT}^{C} = Tfun \cdot \mu_{MZ} \cdot C_{MZ} \cdot \frac{{C_{PRT}}^2}{{C_{PRT}}^2 + {g_{PRT}}^2 + \left(C_{TR} \cdot g_{PRT} / g_{TR}\right)^2}$$

$$GRAZ_{PRT}^{N} = Q_{N,PRT}^{C} \cdot GRAZ_{PRT}^{C}$$

$$GRAZ_{PRT}^{P} = Q_{P,PRT}^{C} \cdot GRAZ_{PRT}^{C}$$

(A6.8) Protozoa net growth rate

$$\frac{dC_{PRT}}{dt} = GROW_{PRT}^{C} - RESP_{PRT} - EXCR_{PRT}^{LDOC} - EXCR_{PRT}^{SDOC} - EXCR2_{PRT}^{SDOC} - POM_{PRT}^{C} - GRAZ_{PRT}^{C}$$

$$\frac{dN_{PRT}}{dt} = GROW_{PRT}^{N} - REMI_{PRT}^{N} - EXCR_{PRT}^{LDON} - EXCR_{PRT}^{SDON} - EXCR2_{PRT}^{SDON} - POM_{PRT}^{N} - GRAZ_{PRT}^{N}$$

$$\frac{dP_{PRT}}{dt} = GROW_{PRT}^{P} - REMI_{PRT}^{P} - EXCR_{PRT}^{LDOP} - EXCR_{PRT}^{SDOP} - EXCR2_{PRT}^{SDOP} - POM_{PRT}^{P} - GRAZ_{PRT}^{P}$$

7. Metazoa (MZ) Processes

(A7.1) Gross growth

$$GROW_{MZ}^{C} = GRAZ_{TR}^{C} + GRAZ_{PRT}^{C}$$

$$GROW_{MZ}^{N} = GRAZ_{TR}^{N} + GRAZ_{PRT}^{N}$$

$$GROW_{MZ}^{P} = GRAZ_{TR}^{P} + GRAZ_{PRT}^{P}$$

(A7.2) DOM excretion

(c) Labile DOM excretion

$$EXCR_{MZ}^{LDOC} = f_{ex,MZ} \cdot ex_{MZ} \cdot GROW_{MZ}^{C}$$

$$EXCR_{MZ}^{LDON} = f_{ex,MZ} \cdot ex_{MZ} \cdot GROW_{MZ}^{N}$$

$$EXCR_{MZ}^{LDOP} = f_{ex,MZ} \cdot ex_{MZ} \cdot GROW_{MZ}^{P}$$

(d) Semi-labile DOM excretion

$$EXCR_{MZ}^{SDOC} = \left(1 - f_{ex,MZ}\right) \cdot ex_{MZ} \cdot GROW_{MZ}^{C}$$

$$EXCR_{MZ}^{SDON} = \left(1 - f_{ex,MZ}\right) \cdot ex_{MZ} \cdot GROW_{MZ}^{N} \cdot Q_{N,MZ}^{C} / q_{N,MZ}^{C}$$

$$EXCR_{MZ}^{SDOP} = \left(1 - f_{ex,MZ}\right) \cdot ex_{MZ} \cdot GROW_{MZ}^{P} \cdot Q_{P,MZ}^{C} / q_{P,MZ}^{C}$$

(A7.3) Respiration

$$RESP_{MZ} = resp_{MZ}^{B} \cdot Tfun \cdot C_{MZ} + resp_{MZ}^{A} \cdot GROW_{MZ}^{C}$$

(A7.4) Semi-labile DOM release in order to adjust stoichiometry

$$EXCR2_{MZ}^{SDOC} = ex_{ADJ,MZ} \cdot C_{MZ} \cdot \max\left(0,\ 1 - \frac{Q_{N,MZ}^{C}}{q_{N,MZ}^{C}},\ 1 - \frac{Q_{P,MZ}^{C}}{q_{P,MZ}^{C}}\right)$$

$$EXCR2_{MZ}^{SDON} = 0.5 \cdot EXCR2_{MZ}^{SDOC} \cdot Q_{N,MZ}^{C}$$
$$EXCR2_{MZ}^{SDOP} = 0.5 \cdot EXCR2_{MZ}^{SDOC} \cdot Q_{P,MZ}^{C}$$

(A7.5) Inorganic nutrient remineralization in order to adjust stoichiometry

$$REMI_{MZ}^{N} = remi_{MZ} \cdot \max\left(0,\ N_{MZ} - q_{N,MZ}^{C} \cdot C_{MZ},\ N_{MZ} - q_{N,MZ}^{C} / q_{P,MZ}^{C} \cdot P_{MZ}\right)$$
$$REMI_{MZ}^{P} = remi_{MZ} \cdot \max\left(0,\ P_{MZ} - q_{P,MZ}^{C} \cdot C_{MZ},\ P_{MZ} - q_{P,MZ}^{C} / q_{N,MZ}^{C} \cdot N_{MZ}\right)$$

(A7.6) POM production

$$POM_{MZ}^{C} = pom_{MZ} \cdot GROW_{MZ}^{C}$$
$$POM_{MZ}^{N} = q_{N,POM}^{C} \cdot POM_{MZ}^{C}$$
$$POM_{MZ}^{P} = q_{P,POM}^{C} \cdot POM_{MZ}^{C}$$

(A7.7) Refractory DOM release

$$REFR_{MZ}^{C} = ex_{REFR,MZ} \cdot C_{MZ}$$
$$REFR_{MZ}^{N} = q_{N,REFR}^{C} \cdot REFR_{MZ}^{C}$$
$$REFR_{MZ}^{P} = q_{P,REFR}^{C} \cdot REFR_{MZ}^{C}$$

(A7.8) Removal by higher-level zooplankton

$$REMV_{MZ}^{C} = remv_{MZ} \cdot C_{MZ}^{2}$$
$$REMV_{MZ}^{N} = Q_{N,MZ}^{C} \cdot REMV_{MZ}^{C}$$
$$REMV_{MZ}^{P} = Q_{P,MZ}^{C} \cdot REMV_{MZ}^{C}$$

(A7.9) POM production by higher-level zooplankton

$$POM_{HZ}^{C} = f_{POM,HZ} \cdot REMV_{MZ}^{C}$$
$$POM_{HZ}^{N} = f_{POM,HZ} \cdot REMV_{MZ}^{N}$$
$$POM_{HZ}^{P} = f_{POM,HZ} \cdot REMV_{MZ}^{P}$$

(A7.10) Semi-labile DOM production by higher-level zooplankton

$$EXCR_{HZ}^{SDOC} = f_{SDOM,HZ} \cdot REMV_{MZ}^{C}$$
$$EXCR_{HZ}^{SDON} = f_{SDOM,HZ} \cdot REMV_{MZ}^{N}$$
$$EXCR_{HZ}^{SDOP} = f_{SDOM,HZ} \cdot REMV_{MZ}^{P}$$

(A7.11) Inorganic nutrients regeneration by higher-level zooplankton

$$REMI_{HZ}^{N} = REMV_{MZ}^{N} - POM_{HZ}^{N} - EXCR_{HZ}^{SDON}$$
$$REMI_{HZ}^{P} = REMV_{MZ}^{P} - POM_{HZ}^{P} - EXCR_{HZ}^{SDOP}$$

(A7.12) Protozoa net growth rate

$$\frac{dC_{MZ}}{dt} = GROW_{MZ}^{C} - RESP_{MZ} - EXCR_{MZ}^{LDOC} - EXCR_{MZ}^{SDOC} - EXCR2_{MZ}^{SDOC} - POM_{MZ}^{C} - REFR_{MZ}^{C} - REMV_{MZ}^{C}$$

$$\frac{dN_{MZ}}{dt} = GROW_{MZ}^{N} - REMI_{MZ}^{N} - EXCR_{MZ}^{LDON} - EXCR_{MZ}^{SDON} - EXCR2_{MZ}^{SDON} - POM_{MZ}^{N} - REFR_{MZ}^{N} - REMV_{MZ}^{N}$$

$$\frac{dP_{MZ}}{dt} = GROW_{MZ}^{P} - REMI_{MZ}^{P} - EXCR_{MZ}^{LDOP} - EXCR_{MZ}^{SDOP} - EXCR2_{MZ}^{SDOP} - POM_{MZ}^{P} - REFR_{MZ}^{P} - REMV_{MZ}^{P}$$

8. Detritus (DET) processes

(A8.1) Dissolution of DET

$$DISS_{DET}^{C} = diss \cdot C_{DET}$$

$$DISS_{DET}^{N} = prf_{N} \cdot diss \cdot N_{DET}$$

$$DISS_{DET}^{P} = prf_{P} \cdot diss \cdot P_{DET}$$

(A8.2) Net Change of DET

$$\frac{dC_{DET}}{dt} = POM_{PHY}^{C} + POM_{TR}^{C} + POM_{UN}^{C} + POM_{PRT}^{C} + POM_{MZ}^{C} + POM_{HZ}^{C} - DISS_{DET}^{C}$$

$$\frac{dN_{DET}}{dt} = POM_{PHY}^{N} + POM_{TR}^{N} + POM_{UN}^{N} + POM_{PRT}^{N} + POM_{MZ}^{N} + POM_{HZ}^{N} - DISS_{DET}^{N}$$

$$\frac{dP_{DET}}{dt} = POM_{PHY}^{P} + POM_{TR}^{P} + POM_{UN}^{P} + POM_{PRT}^{P} + POM_{MZ}^{P} + POM_{HZ}^{P} - DISS_{DET}^{P}$$

9. Inorganic nutrients (NH4,NO3,PO4) processes

(A9.1) Nitrification

$$NTRF = r_{ntrf} \cdot NH4$$

(A9.2) Dissolved inorganic nutrient change rates

$$\frac{dNH4}{dt} = FLUX_{BA}^{NH4} + REMI_{PRT}^{N} + REMI_{MZ}^{N} + REMI_{HZ}^{N} + EXCR_{TR}^{NH4} + EXCR_{UN}^{NH4} - GROW_{PHY}^{NH4} - GROW_{TR}^{NH4} - GROW_{UN}^{NH4} - NTRF$$

$$\frac{dNO3}{dt} = FLUX_{BA}^{NO3} - GROW_{PHY}^{NO3} - GROW_{TR}^{NO3} - GROW_{UN}^{NO3} + NTRF$$

$$\frac{dPO4}{dt} = FLUX_{BA}^{PO4} + REMI_{PRT}^{P} + REMI_{MZ}^{P} + REMI_{HZ}^{P} - GROW_{PHY}^{PO4} - GROW_{TR}^{PO4} - GROW_{UN}^{PO4}$$

10. Dissolved Organic Matter (DOM) processes

(A10.1) Conversion of semi-labile DOM to refractory DOM

$$REFR^{C}_{SDOM} = ex_{REFR,SDOM} \cdot C_{SDOM} \cdot EXP\left(1 - \min\left(\frac{Q^{C}_{N,SDOM}}{q^{C}_{N,REFR}},\ \frac{Q^{C}_{P,SDOM}}{q^{C}_{P,REFR}}\right)\right)$$
$$+ \max\left(0, C_{SDOM} - \frac{N_{SDOM}}{q^{C}_{N.REFR}},\ C_{SDOM} - \frac{P_{SDOM}}{q^{C}_{P.REFR}}\right)$$
$$REFR^{N}_{SDOM} = q^{C}_{N,REFR} \cdot REFR^{C}_{SDOM}$$
$$REFR^{P}_{SDOM} = q^{C}_{P,REFR} \cdot REFR^{C}_{SDOM}$$

(A10.2) DOM change rate

$$\frac{dC_{LDOM}}{dt} = EXCR^{LDOC}_{PHY} + EXCR^{LDOC}_{TR} + EXCR^{LDOC}_{UN} + EXCR^{LDOC}_{PRT} + EXCR^{LDOC}_{MZ} + MORT^{C}_{BA} - GROW^{LDOC}_{BA}$$

$$\frac{dN_{LDOM}}{dt} = EXCR^{LDON}_{PHY} + EXCR^{LDON}_{TR} + EXCR^{LDON}_{UN} + EXCR^{LDON}_{PRT} + EXCR^{LDON}_{MZ} + MORT^{N}_{BA} - GROW^{LDON}_{BA}$$

$$\frac{dP_{LDOM}}{dt} = EXCR^{LDOP}_{PHY} + EXCR^{LDOP}_{TR} + EXCR^{LDOP}_{UN} + EXCR^{LDOP}_{PRT} + EXCR^{LDOP}_{MZ} + MORT^{P}_{BA} - GROW^{LDOP}_{BA}$$

$$\frac{dC_{SDOM}}{dt} = EXCR^{SDOC}_{PHY} + EXCR^{SDOC}_{TR} + EXCR^{SDOC}_{UN} + EXCR^{SDOC}_{BA} + EXCR^{SDOC}_{PRT} + EXCR2^{SDOC}_{PRT}$$
$$+ EXCR^{SDOC}_{MZ} + EXCR2^{SDOC}_{MZ} + EXCR^{SDOC}_{HZ} + DISS^{C}_{DET} - REFR^{C}_{SDOM} - GROW^{SDOC}_{BA}$$
$$\frac{dN_{SDOM}}{dt} = EXCR^{SDON}_{PHY} + EXCR^{SDON}_{TR} + EXCR^{SDON}_{UN} + EXCR^{SDON}_{BA} + EXCR^{SDON}_{PRT} + EXCR2^{SDON}_{PRT}$$
$$+ EXCR^{SDON}_{MZ} + EXCR2^{SDON}_{MZ} + EXCR^{SDON}_{HZ} + DISS^{N}_{DET} - REFR^{N}_{SDOM} - GROW^{SDON}_{BA}$$
$$\frac{dP_{SDOM}}{dt} = EXCR^{SDOP}_{PHY} + EXCR^{SDOP}_{TR} + EXCR^{SDOP}_{UN} + EXCR^{SDOP}_{BA} + EXCR^{SDOP}_{PRT} + EXCR2^{SDOP}_{PRT}$$
$$+ EXCR^{SDOP}_{MZ} + EXCR2^{SDOP}_{MZ} + EXCR^{SDOP}_{HZ} + DISS^{P}_{DET} - REFR^{P}_{SDOM} - GROW^{SDOP}_{BA}$$

11. Diagnostic variables used for data assimilation

(A11.1) Primary production

$$PrPr = GROW^{C}_{PHY} + GROW^{C}_{TR} + GROW^{C}_{UN} - RESP_{PHY} - RESP_{TR} - RESP_{UN}$$
$$- EXCR^{LDOC}_{PHY} - EXCR^{LDOC}_{TR} - EXCR^{LDOC}_{UN} - EXCR^{SDOC}_{PHY} - EXCR^{SDOC}_{TR} - EXCR^{SDOC}_{UN}$$

(A11.2) Heterotrophic bacterial production

$$BPr = GROW^{C}_{BA} - RESP_{BA}$$

Appendix B. Adjoint Model Construction

Let the vector $\mathrm{x} = (x_1, x_2, \cdots, x_m)$ denote the model parameters. The model can be thought as a sequence of calculations, starting with parameters, where each quantity calculated is a model dependent variable. Let $\mathrm{z} = (z_1, z_2, \cdots, z_N)$ denote the model dependent variables and $J = z_{N+1}$ to represent the cost function. A model can be rewritten as:

$$\begin{aligned} z_1 &= f_1(\mathrm{x}), \\ z_2 &= f_2(\mathrm{x}, z_1), \\ &\cdots \\ z_N &= f_N(\mathrm{x}, z_1, z_2, \cdots, z_{N-1}), \\ z_{N+1} &= f_{N+1}(\mathrm{x}, z_1, z_2, \cdots, z_N) \end{aligned} \quad , \tag{B1}$$

where the last equation is a cost function which is calculated after all the model equations. The target for generating the adjoint model is to construct codes to calculate the gradient of the cost function $J = z_{N+1}$ with respect to each component of the parameter x. However, as z_{N+1} is not explicitly represented by x, the partial derivative calculation would require extensive application of the chain rule. Lagrange multipliers are introduced to avoid this difficulty:

$$L(\mathrm{x}, \mathrm{z}, \lambda) = z_{N+1} - \lambda_1 (z_1 - f_1(\mathrm{x})) - \sum_{i=2}^{N+1} \lambda_i (z_i - f_i(\mathrm{x}, z_1, z_2, \cdots, z_{i-1})) \quad , \tag{B2}$$

where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_{N+1})$ is the vector of Lagrange multipliers. At a saddle point of the Lagrange function, the partial derivatives of L with respect to x, z and λ vanish simultaneously. With

$$\frac{\partial L}{\partial \lambda_n} = 0, \qquad n = 1, 2, \cdots, N+1 \tag{B3}$$

one obtains the model equation (B1). In other words, at the saddle point of the Lagrange function, the original model is satisfied.

With

$$\frac{\partial L}{\partial z_{N+1}} = 1 - \lambda_{N+1} = 0, \qquad \lambda_{N+1} = 1 . \tag{B4}$$

And

$$\begin{aligned} &\frac{\partial L}{\partial z_n} = -\lambda_n + \lambda_{n+1}\frac{\partial f_{n+1}}{\partial z_n} + +\lambda_{n+2}\frac{\partial f_{n+2}}{\partial z_n}\cdots + \lambda_{N+1}\frac{\partial f_{N+1}}{\partial z_n} = 0, \\ &\lambda_n = \sum_{i=n+1}^{N+1}\frac{\partial f_i}{\partial z_n}\lambda_i, \qquad n = N, N-1, \cdots, 1 \end{aligned} . \tag{B5}$$

Thus starting from equation (B4), all the Lagrange multipliers can be calculated, which is done in the reverse order of the model equations. As we see at the saddle point the original model (18) is satisfied and equation (B2) gives $L = z_{N+1} = J$, so that the gradients of them are equal $\nabla_{\mathrm{X}} J = \nabla_{\mathrm{X}} L$. Because $L = L(\mathrm{x}, \mathrm{z}, \lambda)$ and z and λ depend on x, by the chain rule,

$$\left(\nabla_{\mathrm{X}} J\right)_k = \left(\nabla_{\mathrm{X}} L\right)_k = \frac{\partial L}{\partial x_k} + \sum_{i=1}^{N+1}\left(\frac{\partial L}{\partial z_i}\frac{\partial z_i}{\partial x_k} + \frac{\partial L}{\partial \lambda_i}\frac{\partial \lambda_i}{\partial x_k}\right). \tag{B6}$$

And $\partial L / \partial z_i = \partial L / \partial \lambda_i = 0$, so that

$$\left(\nabla_X J\right)_k = \frac{\partial L}{\partial x_k} = \sum_{i=1}^{N+1} \frac{\partial f_i}{\partial x_k} \lambda_i, \qquad 1 \le k \le m \,, \tag{B7}$$

The equation (B7) is the form used to generate adjoint code. The following is an example to show how to construct adjoint code.

Considering two lines of model code:

$$\begin{aligned} Y &= G\left(X, \cdots\right) \\ Z &= F\left(X, Y, \cdots\right), \end{aligned} \tag{B8}$$

the Lagrange function would be:

$$L = \cdots - \lambda_Y \left(Y - G\left(X, \cdots\right)\right) - \lambda_Z \left(Z - F\left(X, Y, \cdots\right)\right) - \cdots . \tag{B9}$$

The saddle point of L gives

$$\frac{\partial L}{\partial Y} = -\lambda_Y + \lambda_Z \frac{\partial F}{\partial Y} + \cdots = 0, \quad \lambda_Y = \lambda_Z \frac{\partial F}{\partial Y} + \cdots \tag{B10}$$

From equation (B7) we see the λ_Y has to be accumulated for every equation where Y appears on the right side, so a model code can be written as

$$\lambda_Y = \lambda_Y + \lambda_Z \frac{\partial F}{\partial Y} \,, \tag{B11}$$

where λ_Y has to be initialized to zero before its first use.

Similarly for X ,

$$\begin{aligned} \lambda_X &= \lambda_X + \lambda_Z \frac{\partial F}{\partial X} \\ \lambda_X &= \lambda_X + \lambda_Y \frac{\partial G}{\partial X} \,. \end{aligned} \tag{B12}$$

As indicated by equation (B4) and (B5), the procedure of calculating Lagrange multipliers have to be done in reverse order. Thus in equation (B12) λ_X is calculated first from the second line of the model code and then the first line. Also, the method requires each line be differentiable.

Appendix C. Limited-memory Quasi-Newton Method

M1QN3 solves the problem of minimizing an unconstrained function f of n variables:

$$\min\left\{ f\left(x\right) : x \in R^n \right\},$$

(C1)

where the function f is supposed to be smooth (twice-differentiable). In our application f is the cost function and x are model parameters. The method is iterative, starting with an initial point x_0, or initial guess of model parameters in our application, and producing a sequence of points x_k that converge to the optimum x^*. Denote the f at the point x_k by f_k and its gradient by ∇f_k, and the Hessian matrix (the matrix of second partial derivatives) by B_k.

Here we first discuss how to solve the problem with the original Newton's method. The new iterate x_{k+1} is found as a function of x_k as follows. For any point x, define $p = x - x_k$. The second order Taylor approximation around x_k is

$$m_k\left(p\right) = f_k + p^T \nabla f_k + \frac{1}{2} p^T B_k p \,. \tag{C2}$$

This defines a quadratic model of the function near the point x_k. Its gradient with respect to x_k is $\nabla m_k\left(p\right) = \nabla f_k + B_k p$. Thus the equation (C2) is minimized when its gradient equals zero, or

$$p_k = -{B_K}^{-1} \nabla f_k \,. \tag{C3}$$

p_k thus represents the descent direction of the cost function. Here the function is f

assumed to be convex. (However, according to Nocedal & Wright (2006), the convexity assumption can be relaxed, in which case the method can be implemented in such a way that it is still guaranteed to converge to a local minimum.) Therefore, finding p_k requires computing the inverse of Hessian matrix which may be computationally expensive.

The next point x_{k+1} is then found via a line-search in the direction of p_k: for some $\alpha_k \in (0,\infty)$, $x_{k+1} = x_k + \alpha_k p_k$. In M1QN3 α_k is determined to satisfy the following two Wolfe's conditions:

$$f\left(x_k + \alpha_k p_k\right) \le f\left(x_k\right) + \omega_1 \alpha_k \left\langle \nabla f_k, p_k \right\rangle, \tag{C4}$$

$$\left\langle \nabla f\left(x_k + \alpha_k p_k\right), p_k \right\rangle \ge \omega_2 \left\langle \nabla f_k, p_k \right\rangle. \tag{C5}$$

where $\langle \cdot, \cdot \rangle$ represents inner product. It is necessary to have $0 < \omega_1 < 1/2$ and $\omega_1 < \omega_2 < 1$. M1QN3 set $\omega_1 = 0.0001$ and $\omega_2 = 0.9$.

Because the computation of the inverse of Hessian matrix, denoted as $H_k = B_k^{-1}$ at iteration k, may be very expensive, there are some quasi-Newton methods that only compute an approximation of H_k that is based on the change in gradient between iterations $k-1$ and k. M1QN3 uses the BFGS method, which is considered to be the most effective quasi-Newton method to compute H_k. It is more numerically stable and has very effective self-correcting properties. Defining the displacement vector $s_{k-1} = x_k - x_{k-1}$ and the change in gradient $y_{k-1} = \nabla f_k - \nabla f_{k-1}$, and $\rho_{k-1} = \left(y_{k-1}{}^T s_{k-1}\right)^{-1}$, the BFGS method derives the solution as

$$H_k = \left(I - \rho_{k-1} s_{k-1} {y_{k-1}}^T\right) H_{k-1} \left(I - \rho_{k-1} y_{k-1} {s_{k-1}}^T\right) + s_{k-1} \rho_{k-1} s_{k-1}^T, \quad \text{(C6)}$$

where the new inverse Hessian approximation H_k depends on H_{k-1}, s_{k-1} and y_{k-1}, all of which has been computed from previous iteration. Take the new result back to equation (C3) and then the descent direction p_k can be calculated.

For the initial inverse Hessian matrix H_0, M1QN3 uses the identity matrix which in practice is satisfactory.

However, it is not possible to use BFGS on problems with a very high number n of variables because it is impossible to store or manipulate the approximate inverse Hessian matrix H, which is of size n^2. M1QN3 uses a limited-memory BFGS to solve the problem by storing H in a compressed form that requires storing only a constant multiple of vectors of length n. For a model defined integer m, at iteration $k \geq m+1$, $H_k \nabla f_k$ is directly computed by the following m pairs of vectors:

$$\left\{\left(y_i, s_i\right) :\ k-m \leq i \leq k-1\right\}. \quad \text{(C7)}$$

The computed $H_k \nabla f_k$ can be brought back to equation (C3) to compute descent direction directly. Thus only $O(mn)$ elements are stored compared to store the H_k of size $O\left(n^2\right)$. A higher m can make convergence quicker but requires more storage. A default $m = 7$ suggested by M1QN3 works fine. More mathematical details are given in Gilbert & Lemaréchal (1989).

References

Allen AE, Howard-Jones MH, Booth MG, Frischer ME, Verity PG, Bronk DA, Sanderson MP (2002) Importance of heterotrophic bacterial assimilation of ammonium and nitrate in the Barents Sea during summer. J. Mar. Syst. 38:93-108

Amon RMW, Benner R (1994) Rapid cycling of high-molecular weight dissolved organic matter in the ocean. Nature 369:549-552

Armstrong RA (2006) Optimality-based modeling of nitrogen allocation and photo acclimation in photosynthesis. Deep-Sea Res. II 53:513-531

Benitez-Nelson CR, Bidigare RR, Dickey TD, Landry MR, Leonard CL, Brown SL, Nencioli F, Rii YM, Maiti K, Becker JW, Bibby TS, Black W, Cai WJ, Carlson CA, Chen FZ, Kuwahara VS, Mahaffey C, McAndrew PM, Quay PD, Rappe MS, Selph KE, Simmons MP, Yang EJ (2007) Mesoscale eddies drive increased silica export in the subtropical Pacific Ocean. Science 316:1017-1021

Benner R (2002) Chemical composition and reactivity. In: Hansell DA, Carlson CA (eds) Biogeochemistry of marine dissolved organic matter. Academic Press, San Diego, p 59-90

Bertilsson S, Berglund O, Karl DM, Chisholm SW (2003) Elemental composition of marine Prochlorococcus and Synechococcus: Implications for the ecological stoichiometry of the sea. Limnol. Oceanogr. 48:1721-1731

Biddanda B, Benner R (1997) Carbon, nitrogen and carbohydrate fluxes during the production of particulate and dissolved organic matter by marine phytoplankton. Limnol. and Oceanog. 43:506-518

Biersmith A, Benner R (1998) Carbohydrates in phytoplankton and freshly produced

dissolved organic matter. Mar. Chem. 63:133-144

Bjørnsen PK (1988) Phytoplankton exudation of organic matter: Why do healthy cells do it? Limnol. Oceanogr. 33:151-154

Bronk DA (2002) Dynamics of DON. In: Hansell DA, Carlson CA (eds) Biogeochemistry of marine dissolved organic matter. Academic Press, San Diego, p 153-249

Capone DG, Ferrier MD, Carpenter EJ (1994) Amino-Acid Cycling in Colonies of the Planktonic Marine Cyanobacterium Trichodesmium-Thiebautii. Appl. Envir. Microbiol. 60:3989-3995

Capone DG, Zehr JP, Paerl HW, Bergman B, Carpenter EJ (1997) Trichodesmium, a globally significant cyanobacterium. Science 276:1221-1229

Carlson CA (2002) Production and removal processes. In: Hansell DA, Carlson CA (eds) Biogeochemistry of marine dissolved organic matter. Elsevier Science, USA, New York, p 91-151

Carlson CA, Ducklow HW, Hansell DA, Smith WO, Jr. (1998) Organic carbon partitioning during spring phytoplankton blooms in the Ross Sea polynya and the Sargasso Sea. Limnol. Oceanogr. 43:375-386

Carlson CA, Ducklow HW, Michaels AF (1994) Annual flux of dissolved organic carbon from the euphotic zone in the northwestern Sargasso Sea. Nature 371:405-408

Carpenter EJ, Capone DG (1992) Significance of *Trichodesmium* blooms in the marine nitrogen cycle. In: Carpenter EJ, Capone DG, Rueter JG (eds) Marine Pelagic Cyanobacteria: Trichodesmium and other Diazotrophs. Kluwer Academic Publishers, Dordrecht, p 211-217

Carpenter EJ, McCarthy JJ (1975) Nitrogen fixation and uptake of combined nitrogenous

nutrients by Oscillatoria (Trichodesmium) thiebautii in the western Sargasso Sea. Limnol. Oceanogr. 20:389-401

Carpenter EJ, Romans K (1991) Major role of the cyanobacterium Trichodesmium in nutrient cycling in the North Atlantic Ocean. Science 254:1356-1358

Christian JR, Karl DM (1994) Microbial community structure at the U.S.-Joint Global Ocean Flux study station ALOHA: Inverse methods for estimating biochemical indicator ratios. J. Geophys. Res. (C Oceans). 99:14,269-14,276

Christian JR, Lewis MR, Karl DM (1997) Vertical fluxes of carbon, nitrogen, and phosphorus in the North Pacific Subtropical Gyre near Hawaii. J. Geophys. Res. 102:15,667-15,677

Christian JR, Verschell MA, Murtugudde R, Busalacchi AJ, McClain CR (2002) Biogeochemical modelling of the tropical Pacific Ocean. I: Seasonal and interannual variability. Deep-Sea Res. II 49:509-543

del Giorgio PA, Cole JJ (1998) Bacterial Growth Efficiency in Natural Aquatic Systems. Ann. Rev. Ecol. Syst. 29:503-541

Doney SC, Glover DM, Najjar RG (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 Res. II 43:591-624

Droop M (1974) The nutrient status of algal cell in continuous culture. Journal of the Marine Biological Association of the United Kingdom 54:825-855

Ducklow HW (2000) Bacterioplankton production and biomass in the oceans. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley, New York, p 85-120

Elser JJ, Hassett RP (1994) A Stoichiometric Analysis of the Zooplankton-Phytoplankton

Interaction in Marine and Fresh-Water Ecosystems. Nature 370:211-213

Fagerbakke KM, Heldal M, Norland S (1996) Content of carbon, nitrogen, oxygen, sulfur and phosphorus in native aquatic and cultured bacteria. Aquat. Microb. Ecol. 10:15-27

Falcón LI, Carpenter EJ, Cipriano F, Bergman B, Capone DG (2004) N_2 fixation by unicellular bacterioplankton from the Atlantic and Pacific oceans: Phylogeny and in situ rates. Appl. Envir. Microbiol. 70:765-770

Falkowski PG (2000) Rationalizing elemental ratios in unicellular algae. Journal of Phycology 36:3-6

Falkowski PG, Dubinsky Z, Santostefano G (1985) Light-enhanced dark respiration in phytoplankton. In: Internationale Vereinigung fur theoretische und angewandte Limnologie/Proceedings., Vol 22, p 2830-2833

Fogg GE (1966) The extracellular products of algae. Oceanogr. Mar. Biol. Ann. Rev. 4:195-212

Friedrichs MAM, Dusenberry JA, Anderson LA, Armstrong RA, Chai F, Christian JR, Doney SC, Dunne J, Fujii M, Hood R, McGillicuddy DJ, Moore JK, Schartau M, Spitz YH, Wiggert JD (2007) Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. J. Geophys. Res. 112:C08001, Doi 10.1029/2006jc003852

Fu FX, Zhang YH, Bell PRF, Hutchins DA (2005) Phosphate uptake and growth kinetics of Trichodesmium (Cyanobacteria) isolates from the North Atlantic Ocean and the Great Barrier Reef, Australia. Journal of Phycology 41:62-73

Geider RJ, MacIntyre HL, Kana TM (1998) A dynamic regulatory model of

phytoplanktonic acclimation to light, nutrients, and temperature. Limnol. Oceanogr. 43:679-694

Giering R (1999) Tangent linear and Adjoint Model Compiler , Users manual 1.4. http://www.autodiff.com/tamc

Gilbert JC, Lemaréchal C (1989) Some numerical experiments with variable-storage quasi-newton algorithms. Mathematical Programming 45:407-435

Glibert PM, Bronk DA (1994) Release of dissolved organic nitrogen by marine diazotrophic cyanobacteria, Trichodesmium spp. Appl. Envir. Microbiol. 60:3996-4000

Goldman JC, Caron DA, Dennett MR (1987) Regulation of gross growth efficiency and ammonium regeneration in bacteria by substrate C:N ratio. Limnol. Oceanogr. 32:1239-1252

Goldman JC, McCarthy JJ, Peavey DG (1979) Growth rate influence on the chemical composition of phytoplankton in oceanic waters. Nature 279:210-215

Gruber N, Sarmiento JL (1997) Global patterns of marine nitrogen fixation and denitrification. Global Biogeo. Cycles 11:235-266

Hascoët L, Pascual V (2004) TAPENADE 2.1 User's Guide. http://hal.inria.fr/inria-00069880/en/

Hawser SP, Oneil JM, Roman MR, Codd GA (1992) Toxicity of Blooms of the Cyanobacterium-Trichodesmium to Zooplankton. J. Appl. Phycol. 4:79-86

Hood RR, Subramaniam A, May LR, Carpenter EJ, Capone DG (2002) Remote estimation of nitrogen fixation by Trichodesmium. Deep-Sea Res. II 49:123-147

Hundsdorfer W, Trompert RA (1994) Method of Lines and Direct Discretization - a

Comparison for Linear Advection. Applied Numerical Mathematics 13:469-490

Karl D, Michaels A, Bergman B, Capone D, Carpenter E, Letelier R, Lipschitz F, Paerl H, Sigman D, Stal L (2002) Dinitrogen fixation in the world's oceans. Biogeochemistry 57/58:47-98

Karl DM, Dore JE, Lukas R, Michaels AF, Bates NR, Knap A (2001) Building the long-term picture: the U.S. JGOFS time-series programs. Oceanography 14:6-17

Karl DM, Hebel DV, Bjoerkman K, Letelier RM (1998) The role of dissolved organic matter release in the productivity of the oligotrophic North Pacific Ocean. Limnol. Oceanogr. 43:1270-1286

Karl DM, Letelier R, Tupas L, Dore J, Christian J, Hebel D (1997) The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. Nature 388:533-538

Kirchman DL (1994) The uptake of inorganic nutrients by heterotrophic bacteria. Microb. Ecol. 28:255-271

Kirchman DL (2000) Uptake and regeneration of inorganic nutrients by marine heterotrophic bacteria. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, Inc., New York, p 261-288

Kirchman DL, Wheeler PA (1998) Uptake of ammonium and nitrate by heterotrophic bacteria and phytoplankton in the sub-Arctic Pacific. Deep-Sea Res. I 45:347-365

Kirk JTO (1994) Light and photosynthesis in aquatic systems. Cambridge University Press, New York, NY, USA

Lawson LM, Spitz YH, Hofmann EE, Long RB (1995) A data assimilation technique applied to a predator-prey model. Bulletin of Mathematical Biology 57:593-617

Letelier RM, Karl DM (1996) Role of Trichodesmium spp. in the productivity of the subtropical North Pacific Ocean. Mar. Ecol. Prog. Ser. 133:263-273

Letelier RM, Karl DM (1998) Trichodesmium spp. physiology and nutrient fluxes in the North Pacific subtropical gyre. Aquat. Microb. Ecol. 15:265-276

Levitus S (1982) Climatological atlas of the World Ocean. In: NOAA Prefessional Paper, Vol 13. US Department of Commerce, Rockville, MD, p 173

Matear RJ (1995) Parameter optimization and analysis of ecosystem models using simulated annealing: a case study at Station P. J. Mar. Res. 53:571-607

McCarthy MD, Hedges JI, Benner R (1998) Major bacterial contribution to marine dissolved organic nitrogen. Science 281:231-234

Michaels AF (2001) Elemental stoichiometry, new production and nitrogen fixation. Oceanography 14:68-74

Michaels AF, Olson D, Sarmiento JL, Ammerman JW, Fanning K, Jahnke R, Knap AH, Lipschultz F, Prospero JM (1996) Inputs, losses and transformations of nitrogen and phosphorus in the pelagic North Atlantic Ocean. Biogeochemistry 35:181-226

Middelburg JJ, Nieuwenhuize J (2000) Nitrogen uptake by heterotrophic bacteria and phytoplankton in the nitrate-rich Thames estuary. Mar. Ecol. Prog. Ser. 203:13-21

Montoya JP, Houll CM, Zehr JP, Hansen A, Villareal TA, Capone DG (2004) High rates of N_2 fixation by unicellular diazotrophs in the oligotrophic Pacific Ocean. Nature 430:1027-1032

Moore JK, Doney SC, Kleypas JA, Glover DM, Fung IY (2002) An intermediate complexity marine ecosystem model for the global domain. Deep-Sea Res. II 49:403-462

Morán XAG, Gasol JM, Pedros-Alio C, Estrada M (2001) Dissolved and particulate primary production and bacterial production in offshore Antractic waters during Austral summer: coupled or uncoupled? Mar. Ecol. Prog. Ser. 222:25-39

Mulholland MR (2007) The fate of nitrogen fixed by diazotrophs in the ocean. Biogeosciences 4:37-51

Mulholland MR, Bronk DA, Capone DG (2004) Dinitrogen fixation and release of ammonium and dissolved organic nitrogen by Trichodesmium IMS101. Aquat. Microb. Ecol. 37:85-94

Mulholland MR, Capone DG (2001) Stoichiometry of nitrogen and carbon utilization in cultured populations of Trichodesmium IMS101: Implications for growth. Limnol. Oceanogr. 46:436-443

Mulholland MR, Ohki K, Capone DG (1999) Nitrogen utilization and metabolism relative to patterns of N2 fixation in cultures of Trichodesmium NIBB1067. Journal of Phycology 355:977-988

Nagata T (2000) Production mechanisms of dissolved organic matter. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, New York, p 121-152

Nocedal J (1980) Updating quasi-Newton matrices with limited storage. Mathematics of computation 35:773-782

Nocedal J, Wright SJ (2006) Numerical Optimization, Second Edition. Springer, Verlag

Ogawa H, Tanoue E (2003) Dissolved organic matter in oceanic waters. Journal of Oceanography 59:129-147

Ohki K, Zehr JP, Falkowski PG, Fujita Y (1991) Regulation of nitrogen-fixation by different nitrogen sources in the marine non-heterocystous cyanobacterium

Trichodesmium sp. NIBB1067. Arch. Microbiol. 156:335-337

Pinker RT, Laszlo I (1992) Global Distribution of Photosynthetically Active Radiation as Observed from Satellites. Journal of Climate 5:56-65

Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1986) Numerical Recipes: The Art of Scientific Computing. Cambridge University Press, New York

Rosati A, Miyakoda K (1988) A General-Circulation Model for Upper Ocean Simulation. Journal of Physical Oceanography 18:1601-1626

Santschi PH, Guo L, Baskaran M, Trumbore S, Southon J, Bianchi TS, Honeyman B, Cifuentes L (1995) Isotopic evidence for the contemporary origin of high-molecular weight organic matter in oceanic environments. Geochim. Cosmochim. Acta 59:625-631

Siegel DA, McGillicuddy DJ, Fields EA (1999) Mesoscale eddies, satellite altimetry, and new production in the Sargasso Sea. Journal of Geophysical Research-Oceans 104:13359-13379

Sweby PK (1984) High-Resolution Schemes Using Flux Limiters for Hyperbolic Conservation-Laws. Siam Journal on Numerical Analysis 21:995-1011

Tanoue E, Nishiyama S, Kamo M, Tsugita A (1995) Bacterial-Membranes - Possible Source of a Major Dissolved Protein in Seawater. Geochim. Cosmochim. Acta 59:2643-2648

Tupas LM, Popp BN, Karl DM (1994) Dissolved organic carbon in oligotrophic waters: Experiments on sample preservation, storage and analysis. Mar. Chem. 45:207-216

Tziperman E, Thacker WC (1989) An Optimal-Control Adjoint-Equations Approach to

Studying the Oceanic General-Circulation. Journal of Physical Oceanography 19:1471-1485

Vallino JJ, Hopkinson CS, Hobbie JE (1996) Modeling bacterial utilization of dissolved organic matter: Optimization replaces Monod growth kinetics. Limnol. Oceanogr. 41:1591-1609

Villareal TA, Carpenter EJ (2003) Buoyancy regulation and the potential for vertical migration in the oceanic cyanobacterium *Trichodesmium*. Microb. Ecol. 45:1-10

White AE, Spitz YH, Karl DM, Letelier RM (2006) Flexible elemental stoichiometry in Trichodesmium spp. and its ecological implications. Limnol. Oceanogr. 51:1777-1790

Williams RG, Follows MJ (1998) The Ekman transfer of nutrients and maintenance of new production over the North Atlantic. Deep-Sea Res. I 45:461-489

Yool A, Martin AP, Fernandez C, Clark DR (2007) The significance of nitrification for oceanic new production. Nature 447:999-1002

Zehr JP, Waterbury JB, Turner PJ, Montoya JP, Omoregie E, Steward GF, Hansen A, Karl DM (2001) Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. Nature 412:635-638

Figure Caption

Figure 1. Flow diagram of the model showing structure and state variables. Dashed borders for "N_2", "Higher Levels" and "Refractory DOM" indicate they are not explicitly modeled. Several state variables are grouped by dashed rectangles. A flow arrow ending on a grouping rectangle means the flow applies to all the state variables inside the rectangle.

Figure 1

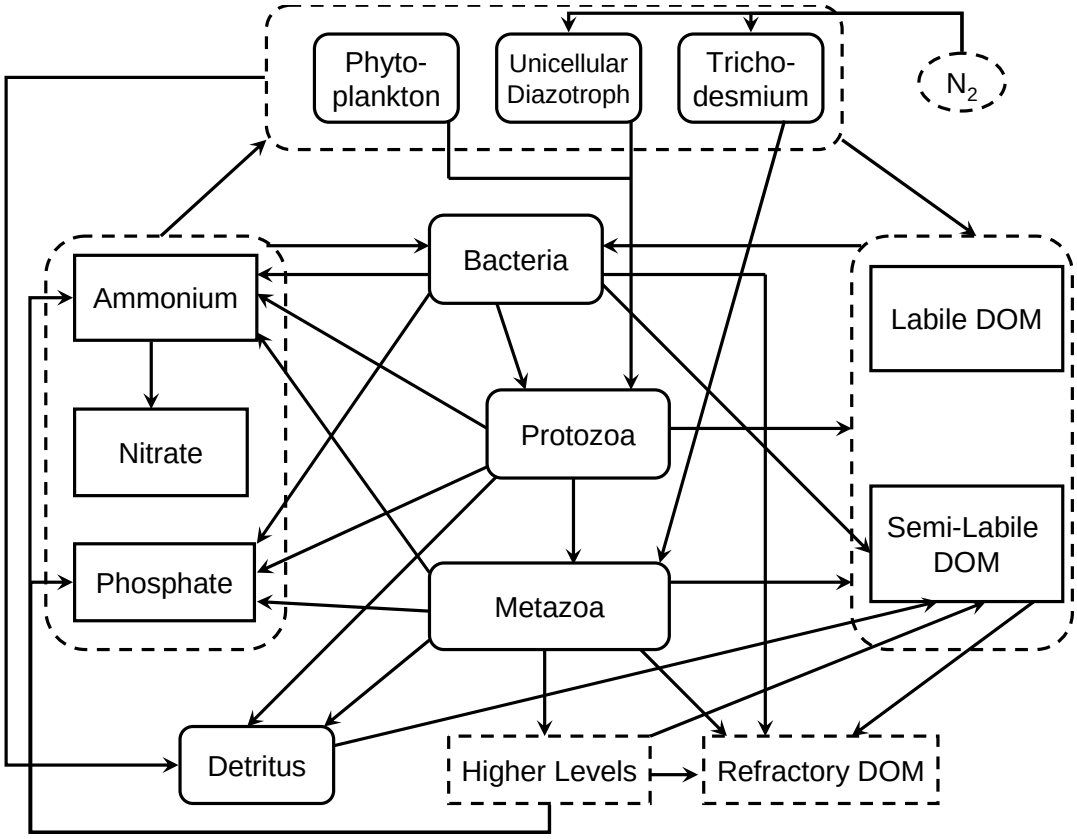

Phyto-
plankton
Unicellular
Diazotroph
Tricho-
desmium
N_2
Bacteria
Ammonium
Labile DOM
Nitrate
Protozoa
Semi-Labile
DOM
Phosphate
Metazoa
Detritus
Higher Levels
Refractory DOM

Table 1. List of the state variables. The tracked components including carbon (C), nitrogen (N), phosphorus (P) and chlorophyll a (CHL) for each state variable are marked with "X".

Name	Component				Definition
	C	N	P	CHL	
PHY	X	X	X	X	Non-diazotrophic phytoplankton
TR	X	X	X	X	Diazotrophic *Trichodesmium* spp.
UN	X	X	X	X	Diazotrophic unicellular phytoplankton
BA	X	X	X		Heterotrophic bacteria
PRT	X	X	X		Protozoan zooplankton
MZ	X	X	X		Metazoan zooplankton
LDOM	X	X	X		Labile dissolved organic matter
SDOM	X	X	X		Semi-labile dissolved organic matter
POM	X	X	X		Particulate organic matter
NH4		X			Ammonium
NO3		X			Nitrate
PO4			X		Phosphate

Table 2. Default model parameter values. The concentrations are in mmol/m^3, and the rates are in day^{-1} unless specified.

Parameter Symbol	Default Value	Units	Description
Following parameters are adjustable:			
A_E	4000		Arrhenius parameter for temperature function
μ_{PHY}	3.5	d^{-1}	C-specific maximum phytoplankton (PHY) growth rate
α_{PHY}^{CHL}	2.0	mol C (g Chl a)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$	Initial slope of Photosynthesis vs. PAR curve of PHY
β_{PHY}	0.002	(W m^{-2})$^{-1}$	Light inhibition parameter for photosynthesis of PHY
$v_{REF,PHY}^{N}$	0.3	mol N (mol C)$^{-1}$ d^{-1}	Maximum nitrogen assimilation rate per PHY carbon biomass
k_{PHY}^{NH4}	0.007	mmol m^{-3}	Ammonium half-saturation concentration for PHY growth
k_{PHY}^{NO3}	0.1	mmol m^{-3}	Nitrate half-saturation concentration for PHY growth
$v_{REF,PHY}^{P}$	0.02	mol P (mol C)$^{-1}$ d^{-1}	Maximum phosphorus assimilation rate per PHY carbon biomass
k_{PHY}^{PO4}	0.005	mmol m^{-3}	Phosphate half-saturation concentration for PHY growth
ζ^{NO3}	2.0	mol C (mol N)$^{-1}$	Carbon requirement (respiration) to assimilate nitrate
θ	4.0	g Chl a (mol N)$^{-1}$	Maximum chlorophyll a to nitrogen ratio
$ex_{PSV,PHY}$	0.05	d^{-1}	PHY passive excretion rate (per biomass)
$ex_{CHO,PHY}$	0.05		PHY carbon hydrate excretion rate (per growth rate)
pom_{PHY}	0.04	(mmol C m^{-3})$^{-1}$$d^{-1}$	POM production rate by PHY aggregation
μ_{TR}	3.0	d^{-1}	C-specific maximum *Trichodesmium* (TR) growth rate
α_{TR}^{CHL}	0.15	mol C (g Chl a)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$	Initial slope of Photosynthesis vs. PAR curve of TR
β_{TR}	0.002	(W m^{-2})$^{-1}$	Light inhibition parameter for photosynthesis of TR
$v_{REF,TR}^{N}$	0.4	mol N (mol C)$^{-1}$ d^{-1}	Maximum nitrogen assimilation rate per TR carbon biomass
k_{TR}^{NH4}	0.05	mmol m^{-3}	Ammonium half-saturation concentration for TR growth
k_{TR}^{NO3}	0.5	mmol m^{-3}	Nitrate half-saturation concentration for TR growth
$v_{REF,TR}^{P}$	0.02	mol P (mol C)$^{-1}$ d^{-1}	Maximum phosphorus assimilation rate per TR carbon biomass
k_{TR}^{PO4}	0.047	mmol m^{-3}	PO4 half-saturation concentration for TR growth

$pick_{TR}^{PO4}$	1.0	d^{-1}	Deep phosphate pick-up (recovery) rate of TR
ζ^{NFIX}	2.5	mol C (mol N)$^{-1}$	Carbon requirement (respiration) to nitrogen fixation
$ex_{PSV,TR}$	0.02	d^{-1}	TR passive excretion rate (per biomass)
$ex_{NFIX,TR}$	0.36		TR DON and ammonium release rate per nitrogen fixation
$ex_{CHO,TR}$	0.05		TR carbon hydrate excretion rate (per growth rate)
pom_{TR}	0.08	(mmol C m^{-3})$^{-1}$ d^{-1}	POM production rate by TR
μ_{UN}	1.6	d^{-1}	C-specific maximum Unicellular N2-fixers (UN) growth rate
α_{UN}^{CHL}	0.5	mol C (g Chl a)$^{-1}$ d^{-1} (W m^{-2})$^{-1}$	Initial slope of Photosynthesis vs. PAR curve of UN
k_{DOM}	0.5	mmol C m^{-3}	Half-saturation concentration of available DOC for heterotrophic bacteria (BA) uptake
$b_{LABI,N}$	5.0		parameter for SDOM's lability related to N content
$b_{LABI,P}$	5.6		parameter for SDOM's lability related to P content
μ_{BA}	2.0	d^{-1}	Maximum BA growth rate
b_{RESP}	0.28	(mmol C m^{-3} d^{-1})$^{-1}$	Parameter control BA active respiration rate versus production
$ex_{ADJ,BA}$	2.0	d^{-1}	Bacteria extra semi-labile DOC excretion rate
$remi_{BA}$	6.0	d^{-1}	Bacteria inorganic nutrients regeneration rate
$ex_{REFR,BA}$	0.018	d^{-1}	Bacteria refractory DOC production rate
$f_{SLCT,BA}$	0.20		BA selection strength on SDOM
$resp_{BA}^{B}$	0.01	d^{-1}	Bacterial basal respiration rate
$r_{\min,BA}^{A}$	0.35		Bacteria minimum active respiration rate
$r_{\max,BA}^{A}$	0.80		Bacteria maximum active respiration rate
$mort_{BA}$	0.021	d^{-1}	Bacteria mortality rate
μ_{PRT}	2.0	d^{-1}	Protozoa (PRT) maximum growth rate
g_{PHY}	0.67	mmol C m^{-3}	Half-saturate density of PHY in PRT grazing function
g_{BA}	1.0	mmol C m^{-3}	Half-saturate density of BA in PRT grazing function
ex_{PRT}	0.20		Total DOM excretion rate per PRT gross growth
$f_{ex,PRT}$	0.75		Fraction of labile DOC of total PRT DOC excretion

$resp^{B}_{PRT}$	0.01	d^{-1}	PRT basal respiration rate
$resp^{A}_{PRT}$	0.67		PRT active respiration rate
$ex_{ADJ,PRT}$	2.0	d^{-1}	PRT extra SDOM excretion rate
$remi_{PRT}$	4.7	d^{-1}	PRT inorganic nutrients regeneration rate
pom_{PRT}	0.027		POM production rate per PRT gross growth
μ_{MZ}	1.3	d^{-1}	Maximum metazoa (MZ) growth rate
g_{PRT}	0.65	mmol C m^{-3}	Half-saturate density of PRT in MZ grazing function
g_{TR}	0.25	mmol C m^{-3}	Half-saturate density of TR in MZ grazing function
ex_{MZ}	0.30		Total DOM excretion rate per MZ gross growth
$f_{ex,MZ}$	0.75		Fraction of labile DOC of total MZ DOC excretion
$resp^{B}_{MZ}$	0.03	d^{-1}	MZ basal respiration
$resp^{A}_{MZ}$	0.25		MZ active respiration
$ex_{ADJ,MZ}$	2.0	d^{-1}	MZ extra semi-labile DOM excretion rate
$remi_{MZ}$	4.0	d^{-1}	MZ inorganic nutrients regeneration rate
pom_{MZ}	0.2		POM production rate per MZ gross growth
$ex_{REFR,MZ}$	0.02	d^{-1}	MZ refractory DOM production rate
$remv_{MZ}$	0.81	$(\text{mmol C } m^{-3})^{-1} d^{-1}$	MZ removal rate by higher-trophic levels
$f_{SDOM,HZ}$	0.2		Fraction of SDOM production by higher-trophic level ingestion
$f_{POM,HZ}$	0.2		Fraction of POM production by higher-trophic level ingestion
$ex_{REFR,SDOM}$	0.0009	d^{-1}	Conversion rate of SDOM to refractory DOM
$q^{C}_{N,REFR}$	0.05	mol N $(\text{mol C})^{-1}$	Refractory DOM N/C ratio
$q^{C}_{P,REFR}$	0.0007	mol P $(\text{mol C})^{-1}$	Refractory DOM P/C ratio
$q^{C}_{N,POM}$	0.12	mol N $(\text{mol C})^{-1}$	N/C ratio for POM generation by PRT and MZ
$q^{C}_{P,POM}$	0.0045	mol P $(\text{mol C})^{-1}$	P/C ratio for POM generation by PRT and MZ
r_{ntrf}	0.1	d^{-1}	Nitrification rate (NH4 to NO3)
prf_{N}	1.5		Preference parameter for dissolving N content in POM
prf_{P}	2.0		Preference parameter for dissolving P content in POM
$wnsv$	5.0	m	Detritus sinking velocity
$diss$	0.02	d^{-1}	POM dissolution rate

Following parameters are assigned to equal values of other parameters			
β_{UN}	β_{PHY}	$(W\ m^{-2})^{-1}$	Light inhibition parameter for photosynthesis of UN
$v^{N}_{REF,UN}$	$v^{N}_{REF,PHY}$	mol N $(mol\ C)^{-1}\ d^{-1}$	Maximum nitrogen assimilation rate per UN carbon biomass
k^{NH4}_{UN}	k^{NH4}_{PHY}	$mmol\ m^{-3}$	NH4 half-saturation concentration for UN growth
k^{NO3}_{UN}	k^{NO3}_{PHY}	$mmol\ m^{-3}$	NO3 half-saturation concentration for UN growth
$v^{P}_{REF,UN}$	$v^{P}_{REF,PHY}$	mol P $(mol\ C)^{-1}\ d^{-1}$	Maximum phosphorus assimilation rate per UN carbon biomass
k^{PO4}_{UN}	k^{PO4}_{PHY}	$mmol\ m^{-3}$	Phosphate half-saturation concentration for UN growth
$ex_{PSV,UN}$	$ex_{PSV,PHY}$	d^{-1}	UN passive excretion rate (per biomass)
$ex_{CHD,UN}$	$ex_{CHD,PHY}$		UN carbon hydrate excretion rate (per growth rate)
$ex_{NFIX,UN}$	$ex_{NFIX,TR}$		UN DON and ammonium release rate per nitrogen fixation
pom_{UN}	pom_{PHY}	$(mmol\ C\ m^{-3})^{-1}d^{-1}$	POM production rate by UN aggregation
g_{UN}	g_{PHY}		PRT grazing parameter on UN
Following parameters are assigned fixed values			
$q^{C}_{N,\min,PHY}$	0.034	mol N $(mol\ C)^{-1}$	Minimum N/C ratio of PHY
$q^{C}_{P,\min,PHY}$	0.00188	mol P $(mol\ C)^{-1}$	Minimum P/C ratio of PHY
$q^{C}_{N,rdf,PHY}$	0.15	mol N $(mol\ C)^{-1}$	Common (Redfield) N/C ratio of PHY
$q^{C}_{P,rdf,PHY}$	0.0094	mol P $(mol\ C)^{-1}$	Common (Redfield) P/C ratio of PHY
$q^{C}_{N,\max,PHY}$	0.17	mol N $(mol\ C)^{-1}$	Maximum N/C ratio of PHY
$q^{C}_{P,\max,PHY}$	0.0169	mol P $(mol\ C)^{-1}$	Maximum P/C ratio of PHY
$q^{C}_{N,\min,TR}$	0.12	mol N $(mol\ C)^{-1}$	Minimum N/C ratio of TR
$q^{C}_{P,\min,TR}$	0.001	mol P $(mol\ C)^{-1}$	Minimum P/C ratio of TR
$q^{C}_{N,rdf,TR}$	0.16	mol N $(mol\ C)^{-1}$	Common N/C ratio of TR
$q^{C}_{P,rdf,TR}$	0.0035	mol P $(mol\ C)^{-1}$	Common P/C ratio of TR
$q^{C}_{N,\max,TR}$	0.2	mol N $(mol\ C)^{-1}$	Maximum N/C ratio of TR
$q^{C}_{P,\max,TR}$	0.006	mol P $(mol\ C)^{-1}$	Maximum P/C ratio of TR
$q^{C}_{N,\min,UN}$	0.12	mol N $(mol\ C)^{-1}$	Minimum N/C ratio of UN
$q^{C}_{P,\min,UN}$	0.001	mol P $(mol\ C)^{-1}$	Minimum P/C ratio of UN
$q^{C}_{N,rdf,UN}$	0.16	mol N $(mol\ C)^{-1}$	Common ratio of UN

$q^C_{P,rdf,UN}$	0.0035	mol P (mol C)$^{-1}$	Common ratio of UN
$q^C_{N,\max,UN}$	0.2	mol N (mol C)$^{-1}$	Maximum N/C ratio of UN
$q^C_{N,\max,UN}$	0.006	mol P (mol C)$^{-1}$	Maximum P/C ratio of UN
$q^C_{N,BA}$	0.18	mol N (mol C)$^{-1}$	Common N/C ratio of BA
$q^C_{P,BA}$	0.02	mol P (mol C)$^{-1}$	Common P/C ratio of BA
$q^C_{N,PRT}$	0.2	mol N (mol C)$^{-1}$	Common N/C ratio of PRT
$q^C_{P,PRT}$	0.022	mol P (mol C)$^{-1}$	Common P/C ratio of PRT
$q^C_{N,MZ}$	0.2	mol N (mol C)$^{-1}$	Common N/C ratio of MZ
$q^C_{P,MZ}$	0.008	mol P (mol C)$^{-1}$	Common P/C ratio of MZ

Table 3. Assimilative data and their availability for each year. “X” represents the data are available, “D” represents data that are not available but derived from other data, and empty boxes represent the data are not available and kept empty.

	NO3	PO4	PHY	CHL	PrPr	Z	BAC	BPr	sDOC	sDON	sDOP	POC	PON	POP	STc	STn	STp
HOT 1990	Y	Y	Y	Y	Y	D	Y		D	Y	Y	Y	Y	Y	Y	Y	Y
HOT 1991	Y	Y	Y	Y	Y	D	Y		D	Y	Y	Y	Y	Y	Y	Y	Y
HOT 1992	Y	Y	Y	Y	Y	D	Y		D	Y	Y	Y	Y	Y	Y	Y	Y
HOT 1993	Y	Y	Y	Y	Y	D	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y
HOT 1994	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y
HOT 1995	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y
HOT 1996	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y
HOT 1997	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y
HOT 1998	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y
HOT 1999	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y
HOT 2000	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
HOT 2001	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
HOT 2002	Y	Y	Y	Y	Y	D	Y	Y	Y	D	D	Y	Y	Y	Y	Y	Y
HOT 2003	Y	Y	Y	Y	Y	D	Y		Y	D	D	Y	Y	Y	Y	Y	Y
HOT 2004	Y	Y	Y	Y	Y	D	Y	Y	Y	D	D	Y	Y	Y	Y	Y	Y
HOT 2005	Y	Y	Y	Y	Y	D	Y	Y	Y	D	D	Y	Y	Y	Y	Y	Y
AS	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y		Y	Y		Y	Y	
EQP	Y	Y	Y	Y	Y	Y	Y	Y	Y			Y	Y		Y	Y	

NO3 & PO4: nitrate and phosphate; **PHY**: Phytoplankton N biomass; **CHL**: Chlorophyll a; **PrPr**: Primary Production; **Z**: Mesozooplankton C biomass; **BAC**: heterotrophic bacterial C biomass; **BPr**: heterotrophic bacterial production; **sDOC, sDON & sDOP**: semi-labile DOM; **STc, STn & STp**: 150m sediment trap of C, N and P flux.

Chapter 3. Heterotrophic Microbial Dynamics in the Open Ocean Revealed by Data Assimilative Modeling

Abstract

Heterotrophic bacteria are major players in ocean carbon and nutrient cycles. To study marine microbial dynamics and their role in open ocean ecosystems, we constructed a one-dimensional epipelagic ecosystem model and applied it to three open ocean sites including the Arabian Sea, the equatorial Pacific and Sta. ALOHA near Hawaii. The model tracks carbon, nitrogen and phosphorus with flexible stoichiometry and includes a detailed scheme for heterotrophic bacteria and dissolved organic matter (DOM). Data assimilation technology was successfully applied to optimize the model parameters and resulted in simultaneously fitting of observed nitrate, phosphate, phytoplankton and zooplankton biomass, primary production, heterotrophic bacterial biomass and production, DOM, and suspended and sinking particulate organic matter. Across the three different study sites, heterotrophic bacteria similarly process nearly half of primary production (PP) and recycle nutrients equivalent to a quarter of the total requirement by primary producers. Semilabile DOM supports variable 11 – 37% of heterotrophic bacterial carbon demand in the three sites, and appears as a temporally-steady carbon resource compared to labile DOM. Following observations, bacterial growth efficiency was parameterized to vary with bacterial carbon demand in the model, which occasionally resulted in amplified bacterial production (BP) variations with respect to PP. Low primary productivity in the oligotrophic Sta. ALOHA may inhibit the variability of bacterial growth efficiency and BP. Our model estimates that roughly 60% of BP is consumed by protists, thus supporting the microbial loop, with most of the remaining BP recycled to labile DOM through viral mortality or released as refractory DOM. In addition to sinking detritus and physical transport of DOM, newly produced

refractory DOM is another form of organic matter export. Our model shows that it contributes between a third to a half of total carbon export, and up to a quarter of total nitrogen export, at the Arabian Sea and Sta. ALOHA; although its contribution is much lower at equatorial Pacific because of the dominant physical export.

1. Introduction

Organic carbon is exported from the upper ocean as both particulate (POM) (Berelson 2001, Honjo et al. 2008) and dissolved organic matter (DOM) (Carlson et al. 1994, Ducklow et al. 1995, Hansell 2002). DOM is potentially an important form of export; it is the dominant organic carbon pool in the ocean (Hedges 1992), and about 50% of primary production by phytoplankton in oceanic ecosystems is ultimately released as DOM through a variety of processes (Ducklow & Carlson 1992, Williams 2000). DOM is mostly utilized by heterotrophic *Bacteria* and possibly also *Archaea* ("bacteria" is used to represent these two groups hereafter) with most organic carbon respired to CO_2 (del Giorgio & Cole 1998, Robinson 2008) and the rest possibly consumed by upper trophic levels. In addition, bacterial activity also regenerates inorganic nutrients from DOM, which in turn directly impacts primary production. Thus bacteria are a key component of the oceanic carbon cycle.

The marine DOM pool represents a broad continuum in terms of its biological reactivity (Carlson 2002), from refractory material turning over on time scales of centuries to labile material turning over on time scales of minutes to days, and semilabile material with turnover times of months to a few years. With a much longer time scale than biological processes, refractory DOM can be considered to be bio-inactive and therefore its production can be regarded as another form of carbon export.

It is not clear that sufficient understanding and data have been obtained to create detailed models of bacterial processes (Anderson 2005). Most modern marine ecosystem models still do not have an explicit representation for heterotrophic bacteria and DOM dynamics, except for a few exceptions with simplified processes (e.g. Spitz et al. 2001,

Anderson & Pondaven 2003, Mongin et al. 2003). However, we believe our current understanding of the microbial loop (Azam et al. 1983) is sufficient to construct an explicit microbial component that is compatible with the current complexity of most marine ecosystem models.

In a nitrogen-limited ocean, the new or exported production is mainly supported by vertical nitrate flux from the deep ocean, by N_2 fixation at the surface and possibly by reactive N deposition from the atmosphere (Dugdale & Goering 1967, Duce et al. 2008, Gruber & Galloway 2008). A consensus has been reached that a significant fraction of the nitrogen requirement for new production in the subtropical oceans comes from N_2 fixation (e.g. Michaels et al. 1996, Capone et al. 1997, Gruber & Sarmiento 1997, Karl et al. 1997, Hood et al. 2002, Casciotti et al. 2008). Thus N_2 fixation should be represented in a model in certain areas to correctly simulate primary production and carbon cycling.

Since Redfield (1934) concluded that the elemental composition of seston is relatively stable in the world ocean (Redfield ratio, molar ratio of C:N:P = 106:16:1), marine biogeochemical models have tended to use one of these elements as currency (usually N) by assuming that the other elements are proportional to this currency at the Redfield ratio. However, we argue that a flexible stoichiometry is crucial to represent heterotrophic and diazotrophic microbial dynamics. Elemental composition of heterotrophic bacteria differs markedly from phytoplankton (Goldman et al. 1987) and N_2-fixers tend to require less P when fixing N_2 (Letelier & Karl 1998). DOM appears to be depleted in N and P and with large temporal and spatial variability (Benner 2002), presenting major challenges for modeling bacterial dynamics. Even for phytoplankton, the C:N or C:P ratios tend to exceed the Redfield ratio, and photosynthesis does not cease

when corresponding nutrients are limited (Geider et al. 1998, Bertilsson et al. 2003). This condition also determines the rate and composition of DOM production by phytoplankton (Fogg 1966, Bjørnsen 1988, Nagata 2000) as phytoplankton may release more DOM when photosynthesis is nutrient-limited.

As more observations become available, it becomes a major challenge to combine these numerous data with marine ecosystem models. One solution is to conduct data assimilation to minimize the differences between modeling results and observations by objectively optimizing model parameters (Lawson et al. 1995, Vallino 2000, Spitz et al. 2001, Friedrichs et al. 2007). As field measurements are expensive and/or difficult, usually only a small number of the state variables and rates in a given model are directly measured. Data assimilation can use the available observations to improve confidence in other unknown variables and rates.

We constructed a marine ecosystem model to address these issues. The model emphasizes detailed heterotrophic microbial dynamics and tracks carbon (C), nitrogen (N) and phosphorus (P) of the state variables with flexible stoichiometry. The model was then applied to a 1-dimensional (1-D) framework at three open ocean sites, one each in the North Pacific Subtropical Gyre, Arabian Sea and Equatorial Pacific. Data assimilation was conducted to optimize the modeling results. We investigated the model results in order to study (1) the patterns of DOM and heterotrophic bacterial production and mechanisms that control them, (2) the importance of heterotrophic microbial dynamics in the ocean carbon cycle and (3) the influence of heterotrophic microbial dynamics on nutrient cycling including regeneration, import and export. The comparison of the three study sites in different open ocean regions is used to reveal general patterns in the open

ocean controlled by regional characteristics.

2. Methods

2.1. The Ecosystem Model

The ecosystem model simulates the C, N and P cycles through the state variables including generic phytoplankton (PHY), N_2-fixing *Trichodesmium* spp. (TR) and unicellular diazotrophs (UN), heterotrophic bacteria (BA), protozoan (PRT) and metazoan (MZ) zooplankton, labile DOM (LDOM) and semilabile DOM (SDOM), detritus (DET) and inorganic nutrients (ammonium NH4, nitrate NO3 and phosphate PO4) (**Figure 1**). Chlorophyll *a* of PHY, TR and UN is also simulated separately from the biomass. The two N_2-fixing species were disabled for the Arabian Sea and Equatorial Pacific models to reflect the fact that N_2-fixation is unlikely to contribute significantly to new production in these nitrogen-replete regions. Here we briefly describe the model schemes of heterotrophic microbial dynamics and flexible stoichiometry. The full model description and equations can be found in Chapter 2.

2.1.1. Microbial Dynamics

The heterotrophic microbial scheme in this model differs from other models mainly in the following ways.

(1) The DOM is divided into labile and semilabile DOM. Continuously variable lability of semilabile DOM is determined by its N:C or P:C ratio. An exponential scheme is used to quickly lower the turnover rate of semilabile DOM with reduced N:C or P:C ratio. The lability of labile DOM is assumed to be constant.

(2) Bacteria have the capability to select semilabile DOM with higher N:C and P:C ratios. Bacteria consume semilabile DOM and reduce its N:C and P:C ratios, so that the

semilabile DOM pool becomes more refractory with time if no new substrate is produced.

(3) The bacterial growth efficiency (BGE), the ratio of BP to bacterial carbon demand (BCD, equal to BP + bacterial respiration), varies with the BCD. An increasing BCD reduces the portion of assimilated C to be respired and reserves a higher portion for BP, and therefore a higher BGE.

The model has two types of autotrophic N_2-fixing bacteria represented: *Trichodesmium* spp. (TR) and Unicellular N_2-fixers (UN). They first use dissolved inorganic nitrogen (DIN) to satisfy their N demand. When the DIN level is low, they fix N_2 as a supplementary N source to satisfy their demand. The N_2-fixation rate is limited by temperature and cellular stoichiometric status. As N_2-fixation needs energy from respired carbon in the model, its rate is enhanced when the cellular C:N ratio is high, and vice-versa. Also, a high cellular C:N reflects N limitation, which favors N_2-fixation. Thus light can limit N_2-fixation indirectly through primary production. Similarly, the inorganic P level also indirectly influences the N_2-fixation rate through its limitation on primary production.

2.1.2. Flexible Stoichiometry

The model tracks C, N and P and allows flexible stoichiometry of the state variables. It has been discussed above how important a scheme of flexible stoichiometry is for modeling bacterial dynamics. Here we discuss its application to primary production in our model.

We selected the Droop equation (Droop 1974, McCarthy 1980, Droop 1983) and adopted the model of Geider et al. (1998) with some changes for the phytoplankton growth scheme of our model. The primary production is limited by the cellular N:C and

P:C ratios (N quota and P quota). Maximum, common and minimum N and P quotas are assigned to PHY, TR and UN. The Redfield ratio is the common quota used to represent the general status of phytoplankton cells based on recent studies showing that it is robust as a mean value for phytoplankton composition (Falkowski 2000, Michaels 2001). In our model, the primary production reduces when the N or P quota drops below the common quota until ceasing completely at the minimum quota. In contrast, the inorganic nutrient assimilation rate reduces when the N or P quotas increase above the common quota until ceasing completely at the maximum quota.

An illustrative scenario of primary production as described by the model could be: (1) At the beginning of day N/C is elevated and photosynthetic carbon fixation is at a higher rate than N assimilation, which will reduce the N or P quota and therefore gradually reduce the photosynthesis rate until it totally ceases if the quota reaches the minimum allowed value that may or may not occur. Meanwhile reducing the N or P quota may gradually enhance nutrient assimilation. (2) During night, photosynthesis is totally shut down; the low N or P quota generated during day makes inorganic N or P assimilation high and the cellular quota will be restored. The increasing quota will slow down the inorganic nutrient assimilation rate until it stops if the quota reaches the maximum allowed value.

The flexible stoichiometry used in the model allows phytoplankton to take advantage of high light to store more carbon when ambient nutrient levels are depleted and to store more nutrients when light is low (Bertilsson et al. 2003). In addition, when the N or P quota is lower than the Redfield value, the model determines the excretion of semilabile DOM by phytoplankton with a lower N:C or P:C ratio than the cellular quota

(Fogg 1966, Bjørnsen 1988, Nagata 2000). This helps phytoplankton to stabilize their cellular quota towards the Redfield ratio and also determines the amount and lability of a large portion of the semilabile DOM production. Although there is no evidence that the DOM produced specifically during this stoichiometric adjustment comprises semilabile material, we postulate such DOM constitutes a waste biproduct by phytoplankton and thus may not be readily usable by bacteria; we therefore treat this material as semilabile DOM.

2.2. Study Sites and Modeling Framework

One site for this study was the long-term biogeochemical observatory, Station ALOHA (22.75ºN, 158.00ºW), the field site for the Hawaii Ocean Time-series (HOT) program, established in the North Pacific Subtropical Gyre in 1988 (Karl et al. 2001, Doney & Ducklow 2006). The station is 100 km north of Oahu, Hawaii with a bottom depth of ~ 4700 m. The station is within the weak southward return flow of the eastern gyre and horizontal transport appears relatively weak (Karl et al. 2001), and thus represents a good site to apply 1-D ecological modeling.

For comparison, we selected sites with available data and contrasting physical regimes in the Arabian Sea (AS) and Equatorial Pacific (EQP). US JGOFS off-shore site S7 was the selected AS site, located at 16ºN, 62ºE. The EQP site is located at 0ºN, 140ºW. Because seasonal monsoons occur at the AS and strong upwelling occurs across the EQP compared to the strongly stratified condition at HOT, these two sites are expected to show different microbial dynamics. As the framework for these two sites has already been set up by others (Friedrichs et al. 2006, Friedrichs et al. 2007), the comparison study can be

conducted by embedding our ecosystem model into the existing framework, along with some recent improvements, particularly assimilating bacterial and DOM data.

1-D frameworks for the upper ocean were set up for these sites to cover 1995 for the AS site, 1992 for the EQP site and 1990 – 2005 at HOT. In the present study, only a single year (2002) was used for HOT to facilitate the comparison to the other two sites and also because the data coverage of bacterial production is best for this year. At the sites in the AS and the EQP, 20 layers were used to cover the upper 150 m, with 10 surface layers of 5 m each and 10 bottom layers of 10 m each. At HOT, the framework used 25 layers, with 10 surface layers of 5 m each and 15 bottom layers of 10 m each, covering the upper 200 m water column. A deeper bottom was used for HOT because the observed nutrient levels at 200 m were much more stable that those at 150 m.

Sea surface photosynthetically active radiation (PAR), mixed layer depth, water temperature, vertical velocity and vertical eddy diffusivity were set to force the model. The physical forcings for AS and EQP were set up by others (Friedrichs et al. 2007). We set the physical forcings for HOT mostly derived from the HOT project data site (http://hahana.soest.hawaii.edu/hot/hot-dogs/index.html). (1) PAR was estimated from a clear-sky short-wave radiation model and then corrected for cloud by the National Center for Environmental Prediction (NECP) reanalysis short-wave data (http://www.cdc.noaa.gov/). PAR was assumed to be 46% of total short-wave radiation (Pinker & Laszlo 1992, Kirk 1994). (2) Mixed-layer depth was estimated from monthly HOT data by using the criteria of the depth for a temperature change of 0.5^{o} C from the sea surface (Levitus 1982). (3) Water temperature is directly obtained from the HOT data site. (4) Vertical velocity was estimated from HOT potential density data by tracking the

first derivative of the displacement of an isopycnal layer whose average depth over time is 200 m. By linearly reducing vertical velocity to zero at the surface, vertical profiles of vertical velocity were constructed. (5) Vertical eddy diffusivity is not required above the mixed-layer depth because the model mixes all the materials in the mixed layer. Below the mixed-layer depth, vertical eddy diffusivity was estimated by assuming a value of $1.1\times10^{-4}m^{-2}s^{-1}$ at the mixed-layer depth and being exponentially reduced through depth. All the physical forcings were linearly interpolated to vertical grids and time steps.

Horizontal advective divergence of nitrate and phosphate was computed for EQP, but was neglected for AS, as resulted from the scaling analysis between horizontal advection and vertical transport (Friedrichs et al. 2007). The main reason is because the upwelling at EQP is so strong that only a small portion of the upwelled nutrients can be consumed, therefore the nutrients are transported horizontally at surface. At HOT where nitrate and phosphate levels are consistently low at surface, the vertically transported nutrients are usually quickly consumed at surface and horizontal advective divergence is not likely.

More details of physical forcings and initial and boundary conditions are described in Chapter 2.

2.3. Data Assimilation

Data assimilation was conducted to minimize the misfits between observations and modeled results by objectively optimizing the model parameter values. A variational adjoint scheme (Lawson et al. 1995) was used for this process including the following procedures. (1) A cost function was constructed to evaluate the misfits between the observations and modeled results. (2) An adjoint model was constructed for the

ecosystem model using the auto-differential software TAPENADE 2.1 (developed by Institut National de Recherche en Informatique en Automatique, France, INRIA) to compute the gradients of the cost function with respect to the model parameters. (3) A limited-memory quasi-Newton optimization software M1QN3 3.1 (Gilbert & Lemaréchal 1989) was adopted to use the computed gradients from the adjoint model to determine the direction and the optimal step size by which the model parameters need to be modified in order to reduce the cost function. These procedures were conducted iteratively until the preset criteria of low gradients of the cost function with respect to the model parameters were met. Not all the parameters are constrainable by the assimilated data. Only a small set of parameters were optimized based on diagonal elements of the inverse Hessian matrix (details see Chapter 2).

2.3.1. Observations

Observations to be assimilated were collected for AS and EQP online at http://usjgofs.whoi.edu/jg/dir/jgofs/, and for HOT at http://hahana.soest.hawaii.edu/hot/hot-dogs/interface.html. The bacterial production data at HOT were measured during some of HOT cruises (Church et al. 2006), but not publicly available.

We included as many data types as possible: up to 17 types of observations were assimilated for each site (**Table 1**). Unavailable data types were left empty in the table.

The vertical profiles of most data were assimilated. However, because of the extreme difference in nutrient levels in and below the mixed layer, it is difficult for the model to fit the nutrient profiles. In addition, the nutrient level below the mixed layer is largely determined by physical instead of biological processes and also has less influence

on the primary production compared to the nutrient level at surface. Even at HOT where the mixed layer (average ~ 60 m) is shallower than the euphotic zone (average ~ 105 m) and both biological and physical processes can influence nitrate levels below the mixed layer (Church et al. 2009), the observations show that more than half of primary production still occurs inside the mixed layer. Thus only the averaged nitrate and phosphate concentrations in the mixed layer were assimilated. At HOT where the nitrate level is low in the mixed layer, the high-sensitivity low-level nitrate data were used for this study. Zooplankton tows in the surface ocean provide the integrated biomass so only volume-averaged zooplankton biomass was assimilated.

Phytoplankton biomass was estimated from abundance of *Prochlorococcus*, *Synechococcus* and pico-eukaryotes, with total N biomass estimated from conversion factor values of 9.5 femtogram (fg, 10^{-15}g) N cell^{-1} for *Prochlorococcus*, 50 fg N cell^{-1} for *Synechococcus* (Bertilsson et al. 2003) and 400 fg N cell^{-1} for pico-eukaryotes.

The model equivalent to observed PP is the sum of net PP of PHY, UN and TR minus their DOM excretion because the ^{14}C tracer method used for the primary production measurements typically only captures the particulate production. It is possible that some ^{14}C-labeled dissolved organic carbon (DOC) could adsorb onto glass fiber filters, so that the measured PP may also include a portion of dissolved production (Karl et al. 1998). Since any DOC included in the PP measurement is difficult to quantify, we only used the particulate PP as the model equivalent, which may potentially lead to overestimation of particulate PP by the model after data assimilation.

Heterotrophic bacterial biomass was estimated from measured cell numbers with the conversion factor of 10 fg C cell^{-1} (Ducklow 2000). Measured heterotrophic bacterial C

production rates include 4, 40 and 3 vertical profiles of leucine incorporation rates for AS, EQP and HOT respectively. A conversion factor of 1.5 kg C /mol leucine (Ducklow 2000) was used to derive the production rates at all 3 sites. Heterotrophic bacterial production is the only data type of all the assimilated data that is not publicly available (yet) from the HOT data site.

Semilabile DOM was derived by subtracting the refractory DOM from total DOM. The refractory fraction was estimated by assuming all DOM in the deep ocean (>1000 m) is refractory. At HOT, when semilabile dissolved organic C, N or P data were not available, the observed average C:N:P in other years of 221:21:1 was used to reconstruct missing data in order to construct a better constrain for DOM composition.

The sediment trap data were collected at 150 m in HOT, which is inside the model domain. However, the sediment trap data were collected at ~ 800 m in both AS and EQP, below the model domain. Sinking speed and dissolution rate were defined as parameters, which are constrained by POM concentration and sediment trap data through data assimilation, so that the amount and time when detritus reached the traps were projected by the model.

2.3.2. Cost Function

The cost function J is a metric to evaluate the difference between modeled results a and observations $\hat{a}$. Here a lower cost function J value indicates a better fit of the model to observations. J summarizes the average Chi-squared value (χ^2) of each data type:

$$J = \sum_{m=1}^{M} \frac{1}{N_m} \sum_{n=1}^{N_m} \left(\frac{a_{m,n} - \hat{a}_{m,n}}{\sigma_m} \right)^2 = \sum_{m=1}^{M} \overline{\chi_m^2}, \qquad (1)$$

where m and n represent assimilative data type and data points, M and N are the total numbers of assimilative data types and data points; σ is the target error which is defined as

$$\sigma = \overline{\hat{a}_n} \cdot \frac{e}{q}, \tag{2}$$

where q is the adjustment factor which will be discussed later; $\overline{\hat{a}_n}$ is the mean of the observations, and e is fractional error level of the observation. To reasonably estimate e, we assumed all properties in the mixed layer should be completely mixed, and a perfect measurement without any error should give a profile with same data value at every measured depth in the mixed layer. Thus we can use the ratio of standard deviation to mean value (coefficient of variation CV) of each vertical profile in the mixed layer to represent the error level for that measurement. The averaged CV of all the profiles was used as e, and each data type has a different e value for each site. This method does not apply to mesozooplankton and sediment trap data because there is only one data point per profile for them. Their fractional error level was assumed to be 25%.

The adjustment factor q was used to compute the target error by adjusting the fractional error level to e/q (**Table 1**) when considering the following issues. Primary production is not mixed in the mixed layer and therefore its error level was reduced. But the CV of the primary production at EQP was already very low (6.7%), thus no adjustment was made. Additional uncertainties are related to conversion factors for estimating phytoplankton biomass, bacterial biomass and production so that the target errors for these variables were increased accordingly. The adjustment factors for the C/N/P components of DOM, POM and sinking particles (sediment traps) were set to

$(1/3)^{1/2}$ so that the total contribution of these 3 C/N/P components is considered to be one data type. In addition, we observed very low CVs (< 5%) for nitrate, phosphate and DOC in EQP, and they were increased in order to avoid fitting the model to noise.

3. Results

Figure 1 shows the average carbon, nitrogen and phosphorus stocks (in boxes) and fluxes (arrows) over the whole model domain for each site after data assimilation.

3.1. Fit of Modeling Results to Observations

The data assimilation optimized the model parameters and reduced the cost functions, i.e., the differences between the observations and the modeled results. Although not all the parameters are directly constrainable by assimilated observations, the gradients of the cost functions for most parameters are low enough to indicate the model reaches at least local optima. Even if the global optimum was not achieved, our objectives were satisfied with both reasonable model parameter values and small differences between model results and observations. Table 6 and 7 list the cost function values, parameter values and gradients before and after data assimilation, as well as the confidence ranges of the parameter after data assimilation. All the results discussed below are after data assimilation.

The averaged observations and the averaged model equivalents through the whole model domain (both spatially and temporally) are mostly comparable at all the three sites (**Table 1**). Model equivalents were only averaged for time points with observations, so the modeled averages given in Table 1 do not reflect the overall averages of the one-year model simulations. The model overestimates the semilabile DON concentration at HOT.

It could be reduced if a higher selection strength is given to bacteria so that bacteria could consume more semilabile DON. However, as the model only has one parameter to control the selection strength for both semilabile DON and DOP consumption, the misfits of semilabile DOP would be increased in order to reduce the misfits of semilabile DON.

The difference of individual observations and model equivalents are defined as Chi-squared (χ^2) as discussed above. The physical scheme of the model completely mixes every state variable inside the mixed layer at every time step. Therefore when comparing the observations inside the mixed layer with their uniform model equivalents, differences are inevitable. Using the target error σ introduced above (Equation 2), a χ^2 ~ 1 will represent a good modeling result whereby the observations are fit within the expected data error on average. Higher χ^2 is not unusual as no model is perfect. The highest χ^2 are obtained for primary production at all the 3 sites, because low target fractional error levels e/q were used (**Table 1**). As discussed above, we artificially reduced e with a high q for primary production. We did that not only to compensate for the fact that primary production is not uniform in the mixed layer, but also for a bias of weighing more on primary production than other data types. Actually, the modeled error $a_n - \hat{a}_n$ relative to the observed mean $\bar{\hat{a}}$ is on average still acceptable, ~20%, 29% and 27% respectively for AS, EQP and HOT. The model does not simulate phytoplankton biomass well at AS and EQP, and does not simulate bacterial biomass well at HOT.

When the depth-integrated observations and the depth-integrated model equivalents are compared, most of the error levels become more acceptable and lower than the

point-to-point errors (**Table 1**). This result indicates that the most significant errors exist in the vertical structure rather than in the temporal variability, i.e. overall, we are more confident of the modeled average status of the system and the temporal trends than the shapes of vertical profiles.

3.2. Bacterial Production

As primary production (PP) flows through trophic levels, DOM is produced mainly from excretion by phytoplankton and zooplankton and dissolution of detritus. Our modeled results confirm that a significant portion of PP is ultimately consumed by bacteria via these transformations. The total amount of carbon consumed by bacteria, i.e., BCD, in all the 3 sites ranged 34% - 51% of PP (**Table 2**). A large amount of the BCD is respired, resulting in bacterial growth efficiencies (BGE) ranging from 20% at EQP to 34% and 35% at AS and HOT.

BCD is satisfied by both labile (LDOM) and semilabile DOM (SDOM). The model reveals that SDOM is a significant source for bacteria and supports more than 30% of the BCD at AS and EQP and 11% at HOT.

3.3. DOM Production and Turnover Time

If a rough steady state can be expected for DOM, its production rate is more important than its concentration to determine the long-term carbon flux to bacteria. Figure 2 shows the contribution from different sources relative to the total DOM production rate.

LDOM dominates the total DOM production at AS and HOT. The high SDOM production at EQP is due to the high detritus production and dissolution rate. But ~38% of total SDOC production at EQP is physically exported, and only ~60% of total SDOC

production is used by bacteria (**Figure 2**).

Although phytoplankton directly excrete 13% - 22% of primary production as DOM (**Table 3**), Figure 2 shows that it only accounts for ~ 1/3 or even less of total DOM production. Primary production cycles through the food web, and more DOM is produced from zooplankton, viral-induced bacterial lysis and dissolution of detritus.

The model estimates similar turnover times for bacteria and LDOM at all the 3 sites. About 15% - 20% of bacteria biomass is renewed daily. LDOM is produced and consumed quickly, and its entire stock is replaced about once per day. SDOM persists much longer at HOT than AS and EQP. The oligotrophic condition at HOT determines that SDOM has low lability because of its low nutrient content. EQP has the shortest SDOM turnover time because of both the high SDOM lability and the strong physical transport.

3.4. Nitrogen Budget

The nitrogen demand by primary production is met by recycling and new inputs (**Figure 3**). Recycling dominates the supply of nitrogen to primary production in all 3 sites. The f-ratio (ratio of new nitrogen input to total demand by primary production) is 19%, 32% and 9.6% at AS, EQP and HOT, respectively. Physical nitrate import at HOT accounts for 2.7% of the total demand while the remaining 6.9% is supplied by N_2 fixation. The strong upwelling at EQP supplies a large amount of nitrate and pushes up the f-ratio. Bacteria and zooplankton recycle different proportions of nitrogen at the 3 sites based on the modeled trophic structure. Nevertheless, bacteria and zooplankton are all important in nitrogen recycling in all the 3 sites.

Total nitrogen supply from recycling and import must equal the total demand if the

inorganic nitrogen pool (nitrate plus ammonium) is at steady state. However, nitrogen supply is 4% more than the demand at EQP and leads to accumulation of nitrate and ammonium whose concentrations increase 30% and 17% respectively over 1 year. The 1-D physical scheme at EQP may not be good enough to calculate vertical and horizontal transport adequately. But as the inorganic nitrogen concentration is already high and does not limit primary production at EQP in the model, this accumulation does not influence primary production significantly. The nitrogen supply and demand nearly balanced at AS and HOT.

Nitrogen is exported out of the model domain via sinking detritus, physical transport and production of refractory DOM (**Figure 3**). Because SDOM dominates the total organic pool, Figure 3 distinguishes it from transport of the other biological species. Sinking detritus is the major form of the export at the 3 sites. The significant export of SDOM and other biological species at EQP is mainly via horizontal advection. So although it is a loss to the model, it may still stay in the surface ocean in the adjacent domain. The production of refractory DOM is another important form of export and is 7% and 25% of total nitrogen export at AS and HOT respectively.

The total input including nitrate transport and N_2 fixation is slightly larger than the total export at EQP and HOT. At EQP, the extra nitrogen import accumulates mainly in the nitrate pool, which may not influence the system significantly as discussed above. At HOT, the extra nitrogen accumulates in the nitrate pools and slightly in the semilabile DON pool. As discussed above, the average semilabile DON concentration is overestimated at HOT (**Table 1**), and thus the model may not be correctly simulating either the production or the consumption of semilabile DON. The accumulation of nitrate

at HOT occurs only below the mixed layer. However, the nitrate level inside the mixed layer maintains a uniformly low level (0.001 – 0.028 μM) and is comparable to observations (**Table 1**). As discussed above, the nitrate level in the mixed layer is more important than that below the mixed layer in limiting primary production, thus the accumulation of nitrate below the mixed layer is not expected to significantly influence the primary production at HOT.

4. Discussion

4.1. Model Schemes of Microbial Dynamics

In this section, we discuss the new model schemes we introduced in Section 2.1 and their performance in the model.

(a) Interaction between DOM and Bacteria

The model determines the lability of SDOM by its N or P content. There is evidence that the N:C and P:C profiles of DOM decrease from surface to depth (Benner 2002, Ogawa & Tanoue 2003). Because the DOM with low N:C or P:C ratio in the deep ocean is mostly refractory (Ogawa & Tanoue 2003), it is reasonable to postulate that the lability of SDOM increases with its N:C or P:C ratio and vice versa. The vertical pattern of the N:C and P:C ratios also suggests that bacteria can select DOM with higher N or P contents and therefore remineralize N or P quicker than C when DOM transports from surface to depth. However, the lability of LDOM is assumed by the model to be fixed and does not depend on its N:C or P:C ratios. Although the turnover time of LDOM varies from minutes to days, it is not long relative to typical ecological turnover times. Unlike SDOM whose longer turnover time acts to decouple its production partially from its

consumption, the production rate of LDOM is likely closely coupled with consumption. Therefore, LDOM production is more important than its lability in determining the rate of bacterial production.

Figure 4 compares production and consumption rates of both labile DOC (LDOC) and semilabile DOC (SDOC). The production rate curves in Figure 4 summarize all the sources and sinks for LDOC and SDOC, except the bacterial consumption. The production curve includes not only the biological production but also the physical transport and the conversion of SDOC to refractory DOC. The tight coupling between the production and consumption of LDOC indicates that all the produced LDOC is quickly consumed by bacteria in the model, as expected. However, the production and consumption of SDOC is not closely-coupled, and the consumption varies much less than the production. Particularly at HOT, consumption rate of SDOC is very stable due to its long turnover time (**Table 4**). EQP has the shortest SDOM turnover time, so the variation of SDOC consumption is larger than other sites and roughly follows SDOC production. The consumption of SDOC can be limited by both the bacterial requirements and the SDOM dynamics including its concentration and stoichiometry. The tight coupling between the production and consumption of LDOC in all the sites reveals that the bacteria are starving, which indicates that the SDOM consumption is not limited by bacterial requirement. Thus the SDOM dynamics should be the major factor influencing the SDOC consumption. Further analysis shows that the SDOC consumption varies with its N:C or P:C ratio but not its concentration because the concentration is too high to limit bacterial usage (data not shown).

It is important to note that the negative net production of SDOC occurring at HOT in

certain periods (**Figure 4**) is because of downward physical transport. The high concentration of SDOC at HOT sometimes makes the physical transport (although only ranging ±0.4% of total SDOC per day) quite significant compared to the SDOC production. However, the net physical transport of SDOC is relatively small, only ~5% of the total SDOC pool per year.

(b) Variable BGE

BGE has been observed to have a positive correlation with BP in aquatic ecosystems (del Giorgio & Cole 1998). Bacteria are expected to use most assimilated carbon to satisfy metabolic energy costs in a low productivity region, but they need a smaller portion if their carbon assimilation rate is high. Our model includes a scheme allowing variable BGE by setting a relationship between the active respiration rate and the bacterial carbon demand (Chapter 2 Appendix A 5.7).

The data assimilation is able to constrain this scheme by optimizing the maximum respiration rate and gives the range of BGE of 27 - 50% at AS, 14 - 22% at EQP and 32 - 36% at HOT over the whole model domain (**Figure 5**). A surprisingly high BGE was obtained at the HOT low productivity region (also **Table 2**). Since BGE is BP/BCD, the BCD (carbon supply) has to be increased in order to reduce BGE at HOT because the BP has been constrained by the observations. However, simulated BCD is already ~ 36% of PP at HOT (**Table 2**). The BCD can be only increased by a limited amount even after a large change of the model by reducing the carbon flux through protozoan and metazoan zooplankton, which, however, may be inconsistent with the observed metazoan zooplankton biomass level. Another possible explanation for the high BGE is that the BP measurements at HOT were overestimated by the assumed leucine conversion factor,

which is not well known. Here a value of 1.5 kg C per mole leucine incorporated was assumed, but Alonso-Sáez et al. (2008) reported that the conversion factor ranged between 0.9 – 3.6 kg C per mole leucine incorporated. The modeled BGE for EQP is significant lower than it is at the other two sites. Because BCD/PP is lowest in the EQP (**Table 2**), it is unlikely that the low BGE at EQP is caused by an overestimation of BCD. Thus the low BP/PP at EQP is probably the main reason for its low BGE (**Table 2**). But again, with the uncertainties of the leucine conversion factor, the observed low BP cannot be confirmed for EQP (Kirchman et al. 1995). If the real leucine conversion factor for EQP was twice higher than that for AS or HOT, estimated BP could be doubled for EQP and similar BGE could be obtained for all the three sites.

In addition, the BGE variability at HOT was much narrower than the other 2 sites. The model gives a BP range of 0.14 - 19 mg C m^{-3} day^{-1} for AS, 0.025 – 2.4 mg C m^{-3} day^{-1} for EQP and 0.022 ~ 1.4 mg C m^{-3} day^{-1} for HOT (**Figure 5**). The lower end of BP at HOT is comparable to other 2 sites partially because of the stable supply from semilabile DOM, but the higher end for HOT is much smaller because of its overall low primary productivity. Thus the smaller range of BP limits the variability of BGE at HOT.

4.2. Dynamics of Bacterial Activity

Figure 6 shows the normalized anomalies of depth-integrated modeled BP and PP, as well as the corresponding depth-integrated observations, which are normalized with the modeled averages in order to directly compare to the modeled results. We used normalized anomalies instead of anomalies themselves in order to directly compare relative variability between BP and PP. At AS, the modeled BP and PP are comparable to the observations; but we cannot confirm the modeled variations between observations. At

EQP, modeled BP and PP fit the observations reasonably well, except for a period of high BP around day 90 that is missed in the simulations. Again, the high BP and PP during summer cannot be confirmed without available observations. At HOT, the approximately monthly PP observations confirm the modeled variations. Although the modeled BP is comparable to the available three BP observations, all of the available observations were measured in the spring. We cannot tell whether the modeled BP accurately represents reality in the other seasons.

As the ultimate resource for BP, primary production is expected to control BP (Cole et al. 1988). Bacterial production in the open ocean is maintained in a remarkably constant ratio to primary production, averaging about 15 - 20% based on observations across oligotrophic and oceanic HNLC and upwelling and blooming systems (Cole et al. 1988, Ducklow 1999, 2000). Our model estimates average BP/PP ratios of 17% and 13% for AS and HOT, and a lower 6.7% value of EQP (**Table 2**). The lower BP/PP of EQP is forced by the data assimilation process in order to fit the observations: the depth-integrated BP/PP observations are 17.3±8.6% and 14.8±2.0% at AS and HOT, and 10.0±4.7% at EQP.

However, the coupling between BP and PP may not be instantaneous. The variability of modeled BP follows that of PP quite well, but the BP may lag behind PP for several days during some periods (**Figure 6**). As a large portion of DOM is produced by zooplankton (**Figure 2**), time-lags can be expected because the zooplankton activity lags growth of phytoplankton. The variability of BP at AS is larger than that of PP in several periods, due to the simultaneous large variations of BGE. A similar pattern occurs at EQP in day 150 – 210 with high BGE. The BP and PP variations are more tightly coupled at

HOT, because the absolute BP variation is relatively small (**Figure 5**) and the BGE variation is very flat.

The percentage of SDOM-supported BCD (S-BCD) has a negative correlation with BP at AS and HOT (**Figure 6**), reflecting a relatively steady carbon supply for SDOM at these 2 sites. A higher PP along with a higher production of LDOM enhances BP, while reducing the percentage of S-BCD, although the amount of S-BCD may be unchanged. Thus SDOM tends to be a stabilizing factor for BP at AS and HOT. However, this pattern is different at EQP, where the percentage of S-BCD tends to positively correlate with BP. The turnover time of SDOM is shorter at EQP (**Table 4**) where the strong physical exchange dominates the dynamics. Thus the SDOM at EQP was not a steady resource for bacterial requirement (**Figure 4**). Furthermore, the comparison of Figures 4 and 6 shows that the SDOM production at EQP actually follows the primary production. Thus SDOM does not act as a stabilizing factor for BP at EQP.

The percentage of S-BCD at HOT is the lowest of the 3 sites. Firstly, from the view of the source of SDOM, it is mainly because of much lower production of detritus by phytoplankton (**Figure 1** and **Table 3**), which in turn supplies less SDOM through dissolution. At HOT where picophytoplankton dominate and the production is low, low aggregation and detritus production are possible. More importantly, the assimilated POM profiles and sediment trap observations constrain the model and likely provide a realistic scenario of production and dissolution rates for detritus. Secondly, from the view of the sink of SDOM, the lower percentage of S-BCD at HOT is mainly because of the lower lability of SDOM because of the nutrient limitation. In summary, the low rates of SDOM source and sink are balanced with each other. The balance is controlled by the underlying

oligotrophic environment, in which phytoplankton tend to be small and SDOM tends to be inactive because of nutrient limitation.

4.3. Fate of Bacterial Production

Heterotrophic bacteria play a major role in the ocean carbon cycle not only as a sink via organic carbon respiration, but also as a source. Heterotrophic bacteria are the principal agents for the recovery of otherwise "lost" DOC, sending it back to higher trophic levels through protozoan grazing, i.e., the so called "microbial loop" (Azam et al. 1983, Ducklow 1983, Azam 1998). One of the reasons that BP is measured is to evaluate how much of the organic carbon returns to foodwebs (Ducklow 2000) with an underlying assumption that most BP is grazed. However, by introducing two other sinks for BP including the production of refractory DOM and viral mortality, our model reveals that bacterivory does not represent the total sink of BP (**Table 5**). Although currently assimilated observations cannot constrain these processes, other studies support our estimated rates.

Heterotrophic bacteria can be an important source of refractory DOM from their membranes or by their metabolism that transforms labile DOM to refractory DOM (Tanoue et al. 1995, McCarthy et al. 1998, Carlson 2002). The production rate from bacterial membranes may be a function of bacterial biomass, while the transformation rate from labile to refractory DOM may relate to BCD. Our model simplifies the process by transforming a certain percentage of bacterial biomass to refractory DOM per day, 3.1%, 2.0% and 1.8% respectively for AS, EQP and HOT after data assimilation. The scheme results in 7.9%, 3.7% and 5.7% of BCD being transformed to refractory DOM in AS, EQP and HOT respectively (**Figure 1**). Stoderegger and Herndl (1998) observed that

8.6% of ^{14}C-glucose uptake by bacteria was released as capsular material. Others observed that 1.4 – 6% of initial ^{14}C-leucine or ^{14}C-glucose concentration was released as biologically resistant DOM (Brophy & Carlson 1989, Tranvik 1993, Heissenberger 1996). As only a portion of the initial leucine or glucose was taken up, these numbers should be higher in terms of how much of BCD was transformed to refractory DOM. For comparison, the model estimated that refractory DOM production accounts for 12 – 16% of total bacterial production (**Table 5**).

It is generally thought that viruses are responsible for about 10 - 50% of the total bacterial mortality in surface oceans, and even more in environments like hypoxic zones that are inhabitable by many protozoa (Fuhrman & Noble 1995, Steward et al. 1996, Fuhrman 1999). Our model assumes a certain percentage of bacteria die per day due to viruses: 3.1%(AS), 2.0%(EQP) and 4.9% (HOT). As a result, viral mortality is responsible for 19%, 16% and 36% of total mortality (excluding refractory DOM production) respectively at AS, EQP and HOT (**Table 5**), which is within the range of other studies. Our model assumes that the viral mortality returns bacterial biomass to the labile DOM pool, which in turn supports 2.3 - 11% of BCD at the three sites. Overall, the model estimated that the viral mortality accounts for 12 – 32% of total bacterial production (**Table 5**).

Bacterial production rates measured by the commonly used radioisotope-labeled leucine or thymidine precursors theoretically should equal the sum of all the sinks discussed here. Our model predicts that a significant proportion of the BP does not return to the foodweb. BCD accounts for 34-51% of PP, and after bacterial production, BP accounts for 6.7-17% of PP (**Table 2**). However, the carbon returned to food web is only

4.3-12% of PP (**Figure 1**). As the viral mortality recycles organic matter (0.8-4.1% of PP, **Figure 1**), it should be excluded from the "net" bacterial production that is meaningful for the ecosystem. The bacteria-produced refractory DOM represents a loss of carbon (0.8-2.7% of PP, **Figure 1**) and nutrient elements.

4.4. Export in form of refractory DOM

Studies of the export of organic matter into the deep ocean have usually focused on the export physically transported out of the surface ocean such as sinking particles collected by sediment traps, DOM export by deep mixing events and zooplankton vertical migration (Ducklow et al. 2001). We consider that formation of refractory DOM is another possible form for the export. Although refractory DOM still physically resides in the surface ocean after it is produced, it becomes biologically inactive. Its residence time may be 4000 - 6000 years as calculated by average age of $DO^{14}C$ (Williams & Druffel 1987, Druffel et al. 1992), which is longer than the timescale of ocean circulation (~ 1000 yr). Thus it is uniformly distributed throughout the global ocean (Carlson 2002). If refractory DOM is produced mainly in the surface but degraded uniformly in the whole ocean, its production will be far more than its decomposition at the surface ocean. Therefore most of the produced refractory DOM can be regarded as an export. However, it is not clear where the refractory DOM is degraded. Photochemical decomposition of refractory DOM occurs at the ocean surface by direct absorption of ultraviolet light by chromophores in aquatic environments (Amon & Benner 1996, Benner & Biddanda 1998). A modeling study by Anderson and Williams (1999) estimated a relatively high rate of photooxidation of refractory DOC at 0.18 mmol C m^{-2} day^{-1}, which alone would set the residence time of refractory DOC to be ~2300 years. Such results suggest that the

decomposition of refractory DOC should occur mainly at the surface ocean. If this prediction is correct, the production of refractory DOM at the surface ocean would not qualify as export because it is roughly balanced by the decomposition. But without direct support from observations, the rate is still uncertain. Another study revealed that DOC may decrease as much as 29% as water passes the deep path of the global thermohaline "conveyor belt" from the North Atlantic to North Pacific Ocean (Hansell & Carlson 1998), which suggests that a significant amount of decomposition occurs in the deep ocean. Although the decomposition rate seems too low to support biological activities, there is evidence from the measurement of natural ^{14}C in nucleic acids of bacteria that oceanic bacteria do utilize ^{14}C-depleted old carbon (Cherrier et al. 1999). Also, losses of refractory DOM in deep water may be due to adsorption onto particulate material as it sinks through the water column (Druffel et al. 1996). All these lines of evidence suggest that refractory DOM may be degraded photochemically in the surface ocean and also in the deep ocean by other processes. In other words, it is likely that at least a part of the refractory DOM produced at the surface ocean can be treated as export.

Nevertheless, the refractory DOM production rates generated by our model overestimate the global ^{14}C derived residence time. Assuming for simplicity that refractory DOM is only produced in the surface ocean, the model-estimated production rates of 7.8, 1.8 and 3.6 mmol C m^{-2} day^{-1} at AS, EQP and HOT predict the residence time of refractory DOM to be 53, 231 and 116 years using the average refractory DOM concentration of 40 μM (Carlson 2002) and ocean depth of 3800 m (Menard & Smith 1966). In the model, at least 29% of the refractory DOM production is contributed by bacteria (**Figure 1**). As discussed above, the rate of refractory DOM production is

consistent with other studies. Thus even if the refractory DOM production by zooplankton and SDOM conversion is overestimated, the model-estimated residence time for refractory DOM is still much shorter than the ^{14}C derived age of refractory DOM. It is possible that the biologically inactive DOM observed in those studies is different from the refractory DOM stored in the deep ocean and may have a shorter residence time. Another possibility is that the residence time of refractory DOM in the ocean is overestimated. Since the current estimate is based only on the bulk ^{14}C age, it has been pointed out that the flux would increase greatly if the refractory DOM is a composite pool with a spectrum of DOM older and younger than the bulk age (Mantoura & Woodward 1983, Ogawa & Tanoue 2003). More over, other studies of photochemical degradation have reduced the residence time to 500 – 2300 years based on the photochemical removal rates (Kieber et al. 1989, Mopper et al. 1991, Moran & Zepp 1997, Benner & Biddanda 1998). It can be further reduced if the removal in the deep ocean is also included; if this is indeed the case, our estimated refractory DOM production rates may not be significantly higher than the observations.

5. Conclusions

Our model was applied to 3 open ocean sites with different environments to study the marine ecosystems, particularly heterotrophic bacterial and DOM dynamics. Seventeen types of major biogeochemical observations were used to objectively optimize the model parameters using a data assimilation procedure. The optimized model shows reasonable goodness of fit to all types of observations simultaneously, which provides a confident basis to support our analysis of modeled results and conclusions. This study is also an

example of how a numerical model can directly integrate with available data to improve our understanding of a marine ecosystem.

The major patterns of heterotrophic bacterial and DOM processes revealed by the model, including BP, BGE, BP/PP, DOM sources/sinks, and nutrient recycling pathways, are reasonably consistent with previously documented observations and hypotheses. This indicates that our current understanding of marine bacterial ecology may have sufficient internal consistency such that this understanding can be used to construct a quantitative closed budget for DOM, detritus and the microbial loop. With extremely oligotrophic environment, the model revealed that semilabile DOM produced at HOT is much less active than other two sites, and the percentage of semilabile DOM-supported bacterial carbon demand is also much lower at HOT. At EQP where upwelling and horizontal advection are strong and the residence times of materials are short, the model indicates that semilabile DOM does not strongly stabilize heterotrophic bacterial production comparing other two sites. Despite these differences, the modeled patterns at the three sites still reveal a possible generality for the world's open ocean that bacteria are important in carbon cycles and nutrient recycling. Our model also indicates two hypotheses that need further testing by field and experimental studies: (1) semilabile DOM is an important substrate which acts as a temporally-steady food source for bacteria and may stabilize bacterial production and primary production through nutrient recycling; (2) refractory DOM production from bacteria membrane and conversion of semi-labile DOM may be a missed form of carbon and nutrient export and may reduce nutrient recycling efficiency. We also conclude that an explicit representation of bacterial processes in a marine ecosystem model is feasible and necessary to correctly simulate the

system.

References

Alonso-Sáez L, Vázquez-Domínguez E, Cardelús C, Pinhassi J, Sala MM, Lekunberri I, Balagué V, Vila-Costa M, Unrein F, Massana R, Simó R, Gasol JM (2008) Factors controlling the year-round variability in carbon flux through bacteria in a coastal marine system. Ecosystems 11:397-409

Amon RMW, Benner R (1996) Photochemical and microbial consumption of dissolved organic carbon and dissolved oxygen in the Amazon River System. Geochim. et Cosmochim. Acta 60:1783-1792

Anderson TR (2005) Plankton functional type modelling: running before we can walk? J. Plankt. Res. 27:1073-1081

Anderson TR, Pondaven P (2003) Non-redfield carbon and nitrogen cycling in the Sargasso Sea: pelagic imbalances and export flux. Deep-Sea Res. I 50:573-591

Anderson TR, Williams PJl (1999) A one-dimensional model of dissolved organic carbon cycling in the water column incorporating combined biological-photochemical decomposition. Global Biogeo. Cycles 13:337-349

Azam F (1998) Microbial control of oceanic carbon flux: The plot thickens. Science 280:694-696.

Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad F (1983) The ecological role of water-column microbes in the sea. Mar. Ecol. Prog. Ser. 10:257-263

Benner R (2002) Chemical composition and reactivity. In: Hansell DA, Carlson CA (eds) Biogeochemistry of marine dissolved organic matter. Academic Press, San Diego,

p 59-90

Benner R, Biddanda B (1998) Photochemical transformation of surface and deep marine dissolved organic matter: Effects on bacterial growth. Limnol. and Oceanog. 43:1373-1378

Berelson WM (2001) The Flux Of Particulate Organic Carbon Into The Ocean Interior: A Comparison Of Four US JGOFS Regional Studies. Oceanography 14:59-67

Bertilsson S, Berglund O, Karl DM, Chisholm SW (2003) Elemental composition of marine Prochlorococcus and Synechococcus: Implications for the ecological stoichiometry of the sea. Limnol. Oceanogr. 48:1721-1731

Bjørnsen PK (1988) Phytoplankton exudation of organic matter: Why do healthy cells do it? Limnol. Oceanogr. 33:151-154

Brophy JE, Carlson DJ (1989) Production of biologically refractory dissolved organic carbon by natural seawater microbial populations. Deep-Sea Res 36:497-507

Capone DG, Zehr JP, Paerl HW, Bergman B, Carpenter EJ (1997) Trichodesmium, a globally significant cyanobacterium. Science 276:1221-1229

Carlson CA (2002) Production and removal processes. In: Hansell DA, Carlson CA (eds) Biogeochemistry of marine dissolved organic matter. Elsevier Science, USA, New York, p 91-151

Carlson CA, Ducklow HW, Michaels AF (1994) Annual flux of dissolved organic carbon from the euphotic zone in the northwestern Sargasso Sea. Nature 371:405-408

Casciotti KL, Trull TW, Glover DM, Davies D (2008) Constraints on nitrogen cycling at the subtropical North Pacific Station ALOHA from isotopic measurements of nitrate and particulate nitrogen. Deep Sea Research Part II: Topical Studies in

Oceanography 55:1661-1672

Cherrier J, Bauer JE, Druffel ERM, Coffin RB, Chanton JP (1999) Radiocarbon in marine bacteria: Evidence for the ages of assimilated carbon. Limnol. Oceanogr. 44:730-736

Church MJ, Ducklow HW, Letelier RM, Karl DM (2006) Temporal and vertical dynamics in picoplankton photoheterotrophic production in the subtropical North Pacific Ocean. Aquat. Microb. Ecol. 45:41-53

Church MJ, Mahaffey C, Letelier RM, Lukas R, Zehr JP, Karl DM (2009) Nitrogen fixation in the North Pacific Subtropical Gyre: The role of mesoscale physical forcing. Global Biogeo. Cycles in press

Cole JJ, Findlay S, Pace ML (1988) Bacterial production in fresh and saltwater ecosystems: a cross-system overview. Mar. Ecol. Prog. Ser. 43:1-10

del Giorgio PA, Cole JJ (1998) Bacterial Growth Efficiency in Natural Aquatic Systems. Ann. Rev. Ecol. Syst. 29:503-541

Doney SC, Ducklow HW (2006) A decade of synthesis and modeling in the US Joint Global Ocean Flux Study. Deep-Sea Res. II 53:451-458

Droop M (1974) The nutrient status of algal cell in continuous culture. Journal of the Marine Biological Association of the United Kingdom 54:825-855

Droop MR (1983) 25 years of algal growth kinetics: a personal view. Bot. Mar. 26

Druffel ERM, Bauer JE, Williams PM, Griffin S, Wolgast D (1996) Seasonal variability of particulate organic radiocarbon in the Northeast Pacific Ocean. J. Geophys. Res. 101:C9,20,543-20,552

Druffel ERM, Williams PM, Bauer JE, Ertel JR (1992) Cycling of dissolved and

particulate organic matter in the open ocean. J. Geophys. Res. 97:C10,15,639-15,6

Duce RA, LaRoche J, Altieri K, Arrigo KR, Baker AR, Capone DG, Cornell S, Dentener F, Galloway J, Ganeshram RS, Geider RJ, Jickells T, Kuypers MM, Langlois R, Liss PS, Liu SM, Middelburg JJ, Moore CM, Nickovic S, Oschlies A, Pedersen T, Prospero J, Schlitzer R, Seitzinger S, Sorensen LL, Uematsu M, Ulloa O, Voss M, Ward B, Zamora L (2008) Impacts of atmospheric anthropogenic nitrogen on the open ocean. Science 320:893-897

Ducklow HW (1983) Production and fate of bacteria in the oceans. Bioscience 33:494-501

Ducklow HW (1999) The bacterial content of the oceanic euphotic zone. FEMS Microbiol. - Ecol. 30:1-10

Ducklow HW (2000) Bacterioplankton production and biomass in the oceans. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley, New York, p 85-120

Ducklow HW, Carlson CA (1992) Oceanic bacterial productivity. Adv. Microb. Ecol. 12:113-181

Ducklow HW, Carlson CA, Bates NR, Knap AH, Michaels AF (1995) Dissolved organic carbon as a component of the biological pump in the North Atlantic Ocean. Philosophical Transactions of the Royal Society of London B 348:161-167

Ducklow HW, Steinberg DK, Buesseler KO (2001) Upper Ocean Carbon Export and the Biological Pump. Oceanography 14:50-58

Dugdale RC, Goering JJ (1967) Uptake of new and regenerated forms of nitrogen in primary production. Limnol. Oceanogr. 12:196-206

Falkowski PG (2000) Rationalizing elemental ratios in unicellular algae. Journal of

Phycology 36:3-6

Fogg GE (1966) The extracellular products of algae. Oceanogr. Mar. Biol. Ann. Rev. 4:195-212

Friedrichs MAM, Dusenberry JA, Anderson LA, Armstrong RA, Chai F, Christian JR, Doney SC, Dunne J, Fujii M, Hood R, McGillicuddy DJ, Moore JK, Schartau M, Spitz YH, Wiggert JD (2007) Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. J. Geophys. Res. 112:C08001, Doi 10.1029/2006jc003852

Friedrichs MAM, Hood RR, Wiggert JD (2006) Ecosystem model complexity versus physical forcing: Quantification of their relative impact with assimilated Arabian Sea data. Deep-Sea Res. II 53:576-600

Fuhrman JA (1999) Marine viruses and their biogeochemical and ecological effects. Nature 399:541-548

Fuhrman JA, Noble RT (1995) Viruses and protists cause similar bacterial mortality in coastal seawater. Limnol. Oceanogr. 40:1236-1242

Geider RJ, MacIntyre HL, Kana TM (1998) A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. Limnol. Oceanogr. 43:679-694

Gilbert JC, Lemaréchal C (1989) Some numerical experiments with variable-storage quasi-newton algorithms. Mathematical Programming 45:407-435

Goldman JC, Caron DA, Dennett MR (1987) Regulation of gross growth efficiency and ammonium regeneration in bacteria by substrate C:N ratio. Limnol. Oceanogr. 32:1239-1252

Gruber N, Galloway JN (2008) An Earth-system perspective of the global nitrogen cycle. Nature 451:293-296

Gruber N, Sarmiento JL (1997) Global patterns of marine nitrogen fixation and denitrification. Global Biogeo. Cycles 11:235-266

Hansell DA (2002) DOC in the global ocean carbon cycle. In: Hansell DA, Carlson CA (eds) Biogeochemistry of Marine Dissolved Organic Matter. Academic Press, San Diego, p 685-715

Hansell DA, Carlson CA (1998) Deep-ocean gradients in the concentration of dissolved organic carbon. Nature 395:263-266

Hedges JI (1992) Global biogeochemical cycles: progress and problems. Mar. Chem. 39:67-93

Heissenberger A, G.G. Leppard, and G.J. Herndl (1996) Relationship between the intracellular integrity and the morphology of the capsular envelope in attached and free-living marine bacteria. Appl. Environ. Microbiol. 62:4521-4528

Honjo S, Manganini SJ, Krishfield RA, 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 76:217-285

Hood RR, Subramaniam A, May LR, Carpenter EJ, Capone DG (2002) Remote estimation of nitrogen fixation by Trichodesmium. Deep-Sea Res. II 49:123-147

Karl DM, Dore JE, Lukas R, Michaels AF, Bates NR, Knap A (2001) Building the long-term picture: the U.S. JGOFS time-series programs. Oceanography 14:6-17

Karl DM, Hebel DV, Bjoerkman K, Letelier RM (1998) The role of dissolved organic

matter release in the productivity of the oligotrophic North Pacific Ocean. Limnol. Oceanogr. 43:1270-1286

Karl DM, Letelier R, Tupas L, Dore J, Christian J, Hebel D (1997) The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. Nature 388:533-538

Kieber DJ, McDaniel J, Mopper K (1989) Photochemical source of biological substrates in sea water: implications for carbon cycling. Nature 341:637-639

Kirchman DL, Rich JH, Barber RT (1995) Biomass and biomass production of heterotrophic bacteria along 140 degree W in the Equatorial Pacific: Effect of temperature on the microbial loop. Deep-Sea Res. II 42:603-619

Kirk JTO (1994) Light and photosynthesis in aquatic systems. Cambridge University Press, New York, NY, USA

Lawson LM, Spitz YH, Hofmann EE, Long RB (1995) A data assimilation technique applied to a predator-prey model. Bulletin of Mathematical Biology 57:593-617

Letelier RM, Karl DM (1998) Trichodesmium spp. physiology and nutrient fluxes in the North Pacific subtropical gyre. Aquat. Microb. Ecol. 15:265-276

Levitus S (1982) Climatological atlas of the World Ocean. In: NOAA Prefessional Paper, Vol 13. US Department of Commerce, Rockville, MD, p 173

Mantoura RFC, Woodward EMS (1983) Conservative behavior of riverine DOC in the Severn Estuary: chemical and geochemical implications. Geochim. Cosmochim. Acta 7:1293-1309

McCarthy JJ (1980) Nitrogen. In: Morris I (ed) The Physiological Ecology of Phytoplankton. Blackwell, Oxford, p 191-234

McCarthy MD, Hedges JI, Benner R (1998) Major bacterial contribution to marine dissolved organic nitrogen. Science 281:231-234

Menard HW, Smith SM (1966) Hypsometry of ocean basin provinces. Journal of Geophyscial Research 71:4305-4325

Michaels AF (2001) Elemental stoichiometry, new production and nitrogen fixation. Oceanography 14:68-74

Michaels AF, Olson D, Sarmiento JL, Ammerman JW, Fanning K, Jahnke R, Knap AH, Lipschultz F, Prospero JM (1996) Inputs, losses and transformations of nitrogen and phosphorus in the pelagic North Atlantic Ocean. Biogeochemistry 35:181-226

Mongin M, Nelson DM, Pondaven P, Brzezinski MA, Treguer P (2003) Simulation of upper-ocean biogeochemistry with a flexible-composition phytoplankton model: C, N and Si cycling in the western Sargasso Sea. Deep-Sea Res. I 50:1445-1480

Mopper K, Zhou XL, Kieber RJ, Kieber DJ, Sikorski RJ, Jones RD (1991) Photochemical Degradation of Dissolved Organic Carbon and Its Impact on the Oceanic Carbon Cycle. Nature 353:60-62

Moran MA, Zepp RG (1997) Role of photoreactions in the formation of biologically labile compounds from dissolved organic matter. Limnol. and Oceanog. 42:1307-1316

Nagata T (2000) Production mechanisms of dissolved organic matter. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, New York, p 121-152

Ogawa H, Tanoue E (2003) Dissolved organic matter in oceanic waters. Journal of Oceanography 59:129-147

Pinker RT, Laszlo I (1992) Global Distribution of Photosynthetically Active Radiation as

Observed from Satellites. Journal of Climate 5:56-65

Redfield AC (1934) On the proportions of organic derivations in sea water and their relation to the composition of plankton. In: Daniel RJ (ed) James Johnstone Memorial Volume. University Press of Liverpool, p 177-192

Robinson C (2008) Heterotrophic bacterial respiration. In: Kirchman DL (ed) Microbial Ecology of the Oceans (second edition). John Wiley & Sons, Inc., Hoboken, New Jersey, p 299 - 334

Spitz YH, Moisan JR, Abbott MR (2001) Configuring an ecosystem model using data from the Bermuda Atlantic Time Series (BATS). Deep Sea Research Part II: Topical Studies in Oceanography 48:1733-1768

Steward GF, Smith DC, Azam F (1996) Abundance and production of bacteria and viruses in the Bering and Chukchi Seas. Marine ecology progress series. Oldendorf 131:287-300

Stoderegger K, Herndl GJ (1998) Production and release of bacterial capsular material and its subsequent utilization by marine bacterioplankton. Limnol. Oceanogr. 43:877-884

Tanoue E, Nishiyama S, Kamo M, Tsugita A (1995) Bacterial-Membranes - Possible Source of a Major Dissolved Protein in Seawater. Geochim. Cosmochim. Acta 59:2643-2648

Tranvik LJ (1993) Microbial transformation of labile dissolved organic matter into humic-like matter in seawater. Fems Microbiol. Ecol. 12:177-183

Vallino JJ (2000) Improving marine ecosystem models: use of data assimilation and mesocosm experiments. J. Mar. Res. 58:117-164

Williams PJl (2000) Heterotrophic bacteria and the dynamics of dissolved organic material. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, New York, p 153- 200

Williams PM, Druffel ERM (1987) Radiocarbon in dissolved organic carbon in the central north Pacific Ocean. Nature 330:246-248

Table 1. Comparison of observations and modeling results, showing the target fractional error level e/q, the averaged Chi-squared values J and the error of depth-integrated modeled results at AS, EQP and HOT. The table gives averaged values of all the assimilated data points and their model equivalents.

	NO3	**PO4**	**MZc**	**PHYn**	**CHL**	**PrPr**	**BAc**	**BPr**	**sDOC**
AS									
Observed Average	3.41	0.559	0.793	0.477	0.355	1.63	0.741	0.153	25.8
Modeled Average	3.09	0.596	0.634	0.469	0.368	1.54	0.755	0.149	22.5
Fractional Error e	9.3%	2.5%	25%	9.2%	25%	24%	5.9%	11%	7.8%
Adjustment Factor q	1.00	0.25	1.00	0.75	1.00	4.00	0.50	0.50	0.71
Chi-squared Average J	8.9	2.5	1.1	38	4.3	11	6.2	6.9	4.2
Depth-integrated Error	22%	14%	20%	23%	22%	6.9%	12%	29%	19%
EQP									
Observed Average	4.89	0.671	0.225	0.151	0.230	0.850	0.477	0.0695	19.4
Modeled Average	3.92	0.668	0.189	0.154	0.211	0.858	0.374	0.0623	16.8
Fractional Error e	4.5%	4.5%	25%	8.9%	10%	6.7%	18%	18%	9.6%
Adjustment Factor q	0.50	0.50	1.00	0.75	1.00	1.00	0.50	0.50	0.30
Chi-squared Average J	5.4	4.0	5.3	16	12	19	1.1	3.3	8.3
Depth-integrated Error	20%	14%	30%	28%	15%	11%	25%	43%	18%
HOT									
Observed Average	0.0046	0.031	0.184	0.081	0.124	0.379	0.310	0.0464	32.6
Modeled Average	0.0059	0.055	0.166	0.083	0.125	0.372	0.283	0.0463	34.2
Fractional Error e	63%	27%	25%	30%	18%	23%	5.6%	3.1%	7.8%
Adjustment Factor q	1.00	1.00	1.00	1.50	1.00	4.00	0.50	0.20	0.58
Chi-squared Average J	1.0	0.65	1.0	1.0	10	23	15	2.1	1.3
Depth-integrated Error	40%	84%	13%	21%	12%	9.7%	31%	10%	8.4%
	sDON	**sDOP**	**POC**	**PON**	**POP**	**STc**	**STn**	**STp**	
AS									
Observed Average	2.90		4.72	0.920		1.11	0.134		
Modeled Average	3.25		5.10	0.885		1.03	0.135		
Fractional Error e	13%		15%	15%		25%	25%		
Adjustment Factor q	0.71		0.71	0.71		0.71	0.71		
Chi-squared Average J	2.2		2.4	4.9		5.6	5.0		
Depth-integrated Error	17%		14%	27%		57%	56%		
EQP									
Observed Average			1.95	0.486		0.403	0.0553		
Modeled Average			2.25	0.453		0.408	0.0380		
Fractional Error e			13%	7.7%		25%	25%		
Adjustment Factor q			0.71	0.71		0.71	0.71		
Chi-squared Average J			3.4	11		2.8	3.3		
Depth-integrated Error			16%	22%		48%	48%		
HOT									
Observed Average	2.12	0.101	1.52	0.219	0.009	2.66	0.334	0.0094	
Modeled Average	3.03	0.098	1.35	0.223	0.012	2.83	0.382	0.0085	
Fractional Error e	11%	11%	9.0%	7.9%	15%	25%	25%	25%	
Adjustment Factor q	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	
Chi-squared Average J	3.3	1.2	2.5	1.6	2.0	0.72	0.69	0.44	
Depth-integrated Error	43%	11%	15%	12%	31%	30%	28%	30%	

NO3: Nitrate (in mixed layer), mmol m^{-3}; **PO4**: phosphate (in mixed layer), mmol m^{-3}; **MZc**: metazoan zooplankton carbon biomass, mmol C m^{-3}; **PHYn**: phytoplankton nitrogen biomass (exclude *Trichodesmium* spp.), mmol N m^{-3}; **CHL:** Chloropyll *a*, mg m^{-3}; **PrPr**: particulate primary production, mmol C m^{-3} day^{-1}; **BAc**: heterotrophic bacterial carbon biomass, mmol C m^{-3}; **BP**: heterotrophic bacterial production, mmol C m^{-3} day^{-1}; **sDOC, sDON** & **sDOP**: semilabile dissolved organic carbon, nitrogen and phosphrous, mmol m^{-3}; **POC, PON** & **POP**: particulate organic carbon, nitrogen and phosphorus, mmol m^{-3}; **STc, STn** & **STp**: particle carbon, nitrogen and phosphorus flux collected by sediment traps at 800m for AS and EQP and at 150m for HOT, mmol m^{-2} day^{-1}.

Table 2. Averaged modeled results related to heterotrophic bacterial activity at AS, EQP and HOT. Modeled bacterial carbon demand (BCD) and bacterial production (BP) are normalized to modeled primary production (PP). BGE is bacterial growth efficiency. Last row reveals the percentage of BCD that is supported by semi-labile DOM (SDOM).

	AS	EQP	HOT
BCD/PP[1]	51%	34%	37%
BP/PP[1]	17%	6.7%	13%
BGE	34%	20%	35%
BCD by SDOM	33%	37%	11%

[1] PP here includes both particulate PP and dissolved PP.

Table 3. Fate of primary carbon production (including both particulate and dissolved production) from grazing, detritus formation, excretion as LDOM and SDOM and physical transport of primary producer biomass in AS, EQP and HOT.

	AS	EQP	HOT
Graze	59.7%	60.2%	80.2%
Detritus	18.3%	23.7%	2.5%
LDOM	16.3%	9.6%	12.3%
SDOM	5.5%	4.0%	5.0%
Physical Transport	0.2%	2.5%	<0.1%

Table 4. Turnover time (days) of heterotrophic bacteria, LDOM and SDOM at AS, EQP and HOT. Turnover time of bacteria is calculated by the ratio of biomass to bacterial production. Turnover time of DOM is calculated by the ratio of stock to total DOM production.

	AS	EQP	HOT
Heterotrophic Bacteria	5.0	5.9	6.5
Labile DOM	1.2	0.8	1.1
Semilabile DOM	120	85	1016

Table 5. Fate of heterotrophic bacterial production to grazing by protozoa, loss as refractory DOM, viral mortality and direct export of heterotrophic bacterial biomass by physics at AS, EQP and HOT.

	AS	EQP	HOT
Grazing	69%	64%	57%
Refractory DOM	16%	12%	12%
Viral Mortality	16%	12%	32%
Physical Transport	0.2%	13%	<0.1%

Table 6. Cost function values, including total cost and the costs for each assimilate observation type, before and after data assimilation. 3 modeling sites are Arabian Sea (AS), Equatorial Pacific (EQP) and Hawaii Ocean Time-series Station ALOHA (HOT).

Observation Type	AS		EQP		HOT	
	Before	After	Before	After	Before	After
Nitrate	167.95	8.68	16.66	5.38	2.22	1.13
Phosphate	84.13	2.70	7.66	4.00	1.36	0.88
Mesozooplankton C	9.89	1.10	19.74	5.31	19.96	1.22
Phytoplankton N	79.20	37.86	78.18	16.24	1.13	1.05
Chlorophyll *a*	29.58	4.23	83.75	11.88	180.04	10.46
Primary Production	411.60	10.91	278.02	19.29	2041.63	23.05
Bacterial C	259.50	5.76	7.14	1.09	22.89	14.58
Bacterial Production	17.37	7.01	9.23	3.26	14.60	2.13
Semilabile DOC	38.92	4.17	53.94	8.26	167.23	1.19
Semilabile DON	13.27	2.02			323.85	2.39
Semilabile DOP					96.96	1.27
POC	20.16	2.43	9.66	3.40	1703.82	2.41
PON	15.98	4.82	60.86	10.92	2744.97	1.56
POP					83.66	2.29
Sediment Trap C	102.58	5.57	8.75	2.76	93.22	0.73
Sediment Trap N	114.12	4.97	9.06	3.30	145.17	0.71
Sediment Trap P					1.30	0.42
Total	1364.24	102.24	642.65	95.07	7644.01	67.51

Table 7. Default model parameter values, and optimized model parameter values after data assimilation. Also the confidence ranges of optimized parameters (if constrained) and gradients of parameters with respect to the cost function after data assimilation. The "Marks" indicates if the parameter was constrained (CS), adjusted but not constrained (AJ) or not adjusted (empty) by data assimilation. Please see model description for parameter definitions and units.

(a) ARABIAN SEA (AS):

Parameter Symbol	Default Value	Optimized Value	Confidence Range	Gradient	Marks
A_E	4000	4000.0		0.3530	
μ_{PHY}	3.5	6.4211	5.58 – 7.40	-0.1060	CS
α_{PHY}^{CHL}	2.0	0.4322	0.397 – 0.471	-0.2900	CS
β_{PHY}	0.002	0.0020		-0.7980	
$v_{REF,PHY}^{N}$	0.3	0.3000		-1.8800	
k_{PHY}^{NH4}	0.25	0.2500		1.6200	
k_{PHY}^{NO3}	1.0	1.2500		-0.3680	AJ
$v_{REF,PHY}^{P}$	0.02	0.0200		-1.5600	
k_{PHY}^{PO4}	0.06	0.0600		0.1630	
ζ^{NO3}	2.0	2.0000		1.2800	
θ	2.0	0.8799	0.748 – 1.04	-0.1540	CS
$ex_{PSV,PHY}$	0.05	0.0500		2.4900	
$ex_{CHO,PHY}$	0.05	0.0500		-1.0000	
pom_{PHY}	0.04	0.0227	0.0185 – 0.0278	-0.0191	CS
μ_{TR}	3.0	N/A		N/A	
α_{TR}^{CHL}	0.15	N/A		N/A	
β_{TR}	0.002	N/A		N/A	
$v_{REF,TR}^{N}$	0.4	N/A		N/A	
k_{TR}^{NH4}	0.05	N/A		N/A	
k_{TR}^{NO3}	0.5	N/A		N/A	
$v_{REF,TR}^{P}$	0.02	N/A		N/A	
k_{TR}^{PO4}	0.047	N/A		N/A	
$pick_{TR}^{PO4}$	1.0	N/A		N/A	
ζ^{NFIX}	2.5	N/A		N/A	
$ex_{PSV,TR}$	0.02	N/A		N/A	
$ex_{NFIX,TR}$	0.36	N/A		N/A	
$ex_{CHO,TR}$	0.05	N/A		N/A	
pom_{TR}	0.08	N/A		N/A	
μ_{UN}	1.6	N/A		N/A	
α_{UN}^{CHL}	0.5	N/A		N/A	

k_{DOM}	0.5	0.5000		0.6320	
$b_{LABI,N}$	5.0	6.8000		0.0000	AJ
$b_{LABI,P}$	5.6	5.3741	5.08 – 5.69	-0.0571	CS
μ_{BA}	2.0	1.1395	0.984 – 1.32	0.0545	CS
b_{RESP}	0.28	0.2800		0.1660	
$ex_{ADJ,BA}$	2.0	2.0000		0.0000	
$remi_{BA}$	6.0	6.0000		0.5370	
$ex_{REFR,BA}$	0.018	0.0310		1.0500	AJ
$f_{SLCT,BA}$	0.25	0.2500		-0.5520	
$resp^{B}_{BA}$	0.01	0.0100		0.0627	
$r^{A}_{\min,BA}$	0.35	0.3500		-0.1670	
$r^{A}_{\max,BA}$	0.80	0.6948	0.636 – 0.759	-0.0480	CS
$mort_{BA}$	0.021	0.0310		0.3340	AJ
μ_{PRT}	2.0	0.9931	0.758 – 1.30	0.4860	CS
g_{PHY}	0.67	3.0442	2.67 – 3.47	-0.3470	CS
g_{BA}	1.0	2.0196	1.70 – 2.39	0.0452	CS
ex_{PRT}	0.20	0.2000		0.3660	
$f_{ex,PRT}$	0.75	0.7500		-0.1830	
$resp^{B}_{PRT}$	0.01	0.0100		0.0750	
$resp^{A}_{PRT}$	0.67	0.3243	0.217 – 0.485	-0.1820	CS
$ex_{ADJ,PRT}$	2.0	2.0000		-0.0106	
$remi_{PRT}$	4.0	4.0000		-0.1800	
pom_{PRT}	0.027	0.0280		0.1490	AJ
μ_{MZ}	1.3	1.1316	1.02 – 1.26	-0.7600	CS
g_{PRT}	0.65	1.4645	1.13 – 1.90	0.4670	CS
g_{TR}	2.0	2.0000		0.0000	
ex_{MZ}	0.20	0.2000		2.0400	
$f_{ex,MZ}$	0.75	0.7500		-1.2400	
$resp^{B}_{MZ}$	0.03	0.0300		0.3590	
$resp^{A}_{MZ}$	0.25	0.2500		1.5500	
$ex_{ADJ,MZ}$	2.0	2.0000		0.0000	
$remi_{MZ}$	4.0	4.0000		-0.1490	

pom_{MZ}	0.12	0.1200		1.2000	
$ex_{REFR,MZ}$	0.02	0.0200		0.2870	
$remv_{MZ}$	0.81	0.3800		1.9800	AJ
$f_{SDOM,HZ}$	0.2	0.1400		0.5530	AJ
$f_{POM,HZ}$	0.2	0.0940		0.7240	AJ
$ex_{REFR,SDOM}$	0.0009	0.0009		0.2840	
$q^{C}_{N,REFR}$	0.05	0.0500		1.3600	
$q^{C}_{P,REFR}$	0.0007	0.0007		-0.2950	
$q^{C}_{N,POM}$	0.12	0.1200		3.9900	
$q^{C}_{P,POM}$	0.01	0.0100		-2.2900	
r_{ntrf}	0.15	0.1500		-1.0400	
prf_{N}	1.05	1.0500		-1.0100	
prf_{P}	1.10	1.1000		2.3400	
wnsv	30.0	30.0000		-1.8200	
diss	0.1	0.1192	0.112 – 0.147	0.0318	CS

(b) Equatorial Pacific (EQP):

Parameter Symbol	Default Value	Optimized Value	Confidence Range	Gradient	Marks
A_E	4000	2000.0		5.6500	AJ
μ_{PHY}	3.5	3.4436	3.12 – 3.80	0.2910	CS
α_{PHY}^{CHL}	2.0	1.7479	1.52 – 2.01	0.1650	CS
β_{PHY}	0.001	0.0010		-1.8800	
$v_{REF,PHY}^{N}$	0.3	0.3278	0.282 – 0.380	0.2990	CS
k_{PHY}^{NH4}	0.25	0.1900		0.7680	AJ
k_{PHY}^{NO3}	4.0	4.0000		-5.8200	
$v_{REF,PHY}^{P}$	0.02	0.0140		-0.9730	AJ
k_{PHY}^{PO4}	0.1	0.1000		0.0954	
ζ^{NO3}	2.0	2.0000		-0.4200	
θ	2.0	3.4000		-3.5800	AJ
$ex_{PSV,PHY}$	0.05	0.0500		0.8550	
$ex_{CHO,PHY}$	0.05	0.0500		-1.2700	
pom_{PHY}	0.04	0.1300		-0.9440	AJ
μ_{TR}	3.0	N/A		N/A	
α_{TR}^{CHL}	0.15	N/A		N/A	
β_{TR}	0.002	N/A		N/A	
$v_{REF,TR}^{N}$	0.4	N/A		N/A	
k_{TR}^{NH4}	0.05	N/A		N/A	
k_{TR}^{NO3}	0.5	N/A		N/A	
$v_{REF,TR}^{P}$	0.02	N/A		N/A	
k_{TR}^{PO4}	0.047	N/A		N/A	
$pick_{TR}^{PO4}$	1.0	N/A		N/A	
ζ^{NFIX}	2.5	N/A		N/A	
$ex_{PSV,TR}$	0.02	N/A		N/A	
$ex_{NFIX,TR}$	0.36	N/A		N/A	
$ex_{CHO,TR}$	0.05	N/A		N/A	
pom_{TR}	0.08	N/A		N/A	
μ_{UN}	1.6	N/A		N/A	
α_{UN}^{CHL}	0.5	N/A		N/A	

k_{DOM}	0.5	0.5500		0.0126	AJ
$b_{LABI,N}$	5.0	5.2000		0.0000	AJ
$b_{LABI,P}$	5.6	8.2426	6.81 – 9.98	0.6020	CS
μ_{BA}	2.0	2.9224	1.74 – 4.90	-0.2090	CS
b_{RESP}	0.15	0.1500		-0.6530	
$ex_{ADJ,BA}$	2.0	2.0000		0.0001	
$remi_{BA}$	6.0	6.0000		-0.8400	
$ex_{REFR,BA}$	0.02	0.0200		0.1130	
$f_{SLCT,BA}$	0.25	0.2500		-0.3250	
$resp^{B}_{BA}$	0.01	0.0100		0.0565	
$r^{A}_{\min,BA}$	0.35	0.3500		0.4380	
$r^{A}_{\max,BA}$	0.80	0.8755	0.835 – 0.912	1.6800	CS
$mort_{BA}$	0.02	0.0200		-0.0995	
μ_{PRT}	2.0	1.0986	0.944 – 1.28	-0.1560	CS
g_{PHY}	0.67	0.8992	0.846 – 0.956	0.3720	CS
g_{BA}	1.0	1.5737	1.16 – 2.14	-0.2240	CS
ex_{PRT}	0.20	0.2000		-0.8930	
$f_{ex,PRT}$	0.75	0.7500		1.3900	
$resp^{B}_{PRT}$	0.01	0.0100		3.6000	
$resp^{A}_{PRT}$	0.67	0.5819	0.544 – 0.623	0.5720	CS
$ex_{ADJ,PRT}$	2.0	2.0000		0.0000	
$remi_{PRT}$	4.0	4.0000		0.4440	
pom_{PRT}	0.027	0.0260		-0.1080	AJ
μ_{MZ}	1.3	1.2381	0.937 – 1.64	0.0719	CS
g_{PRT}	2.0	1.5984	1.29 – 1.98	-0.2000	CS
g_{TR}	1.0	1.0000		0.0000	
ex_{MZ}	0.20	0.2000		-0.2030	AJ
$f_{ex,MZ}$	0.75	0.7500		-0.2000	
$resp^{B}_{MZ}$	0.03	0.0300		0.6140	
$resp^{A}_{MZ}$	0.25	0.4461	0.379 – 0.525	-0.3880	CS
$ex_{ADJ,MZ}$	2.0	2.0000		0.0000	
$remi_{MZ}$	4.0	4.0000		-0.0769	

pom_{MZ}	0.1	0.1000		-0.2200	
$ex_{REFR,MZ}$	0.02	0.0200		0.0200	
$remv_{MZ}$	0.8	0.2000		0.2640	AJ
$f_{SDOM,HZ}$	0.2	0.2000		-0.0572	
$f_{POM,HZ}$	0.2	0.3500		-0.0713	
$ex_{REFR,SDOM}$	0.0009	0.0009		0.3510	
$q^{C}_{N,REFR}$	0.05	0.0500		1.0400	
$q^{C}_{P,REFR}$	0.0007	0.0007		0.0035	
$q^{C}_{N,POM}$	0.12	0.1200		0.2090	
$q^{C}_{P,POM}$	0.01	0.0100		-0.3930	
r_{ntrf}	0.15	0.1700		-1.2600	AJ
prf_{N}	1.1	1.1000		10.1000	
prf_{P}	1.1	2.0352	1.37 – 3.03	-0.0634	CS
wnsv	30.0	30.0000		0.9050	
diss	0.1	0.1563	0.150 – 1.63	0.1290	CS

(c) Hawaii Ocean Time-series Station ALOHA (HOT):

Parameter Symbol	Default Value	Optimized Value	Confidence Range	Gradient	Marks
A_E	4000	4000.0		0.1700	
μ_{PHY}	3.5	2.2283	1.91 – 2.60	-0.0165	CS
α_{PHY}^{CHL}	2.0	5.1282	4.00 – 6.57	-0.0080	CS
β_{PHY}	0.0005	0.0005		1.8500	
$v_{REF,PHY}^{N}$	0.3	0.2057	0.143 – 0.297	-0.0567	CS
k_{PHY}^{NH4}	0.01	0.0100		1.7300	
k_{PHY}^{NO3}	0.1	0.1000		0.5030	
$v_{REF,PHY}^{P}$	0.02	0.0067	0.0057 – 0.0079	0.1220	CS
k_{PHY}^{PO4}	0.0048	0.0048		0.0010	
ζ^{NO3}	2.0	2.0000		-0.7800	
θ	2.0	2.6623	2.29 – 3.09	-0.0084	CS
$ex_{PSV,PHY}$	0.05	0.0500		2.9300	
$ex_{CHO,PHY}$	0.05	0.0500		1.0400	
pom_{PHY}	0.038	0.0380		2.0500	
μ_{TR}	3.0	3.0000		3.0100	
α_{TR}^{CHL}	0.25	0.2500		-2.0300	
β_{TR}	0.002	0.0020		-1.2100	
$v_{REF,TR}^{N}$	0.4	0.4000		-0.4450	
k_{TR}^{NH4}	0.05	0.0500		0.0977	
k_{TR}^{NO3}	0.5	0.5000		-0.1450	
$v_{REF,TR}^{P}$	0.02	0.0200		-0.2140	
k_{TR}^{PO4}	0.047	0.0470		0.1150	
$pick_{TR}^{PO4}$	1.0	1.0000		0.0282	
ζ^{NFIX}	2.5	2.5000		0.0246	
$ex_{PSV,TR}$	0.02	0.0200		0.6090	
$ex_{NFIX,TR}$	0.36	0.3600		0.0348	
$ex_{CHO,TR}$	0.05	0.0500		0.3200	
pom_{TR}	0.1	0.0800		0.0586	AJ
μ_{UN}	1.6	1.6000		-4.6200	

α_{UN}^{CHL}	0.5	0.5000		-1.3400	
k_{DOM}	0.5	0.5000		0.4890	
$b_{LABI,N}$	10.0	10.0000		-0.0004	
$b_{LABI,P}$	10.0	8.9327	8.68 – 9.19	0.0225	CS
μ_{BA}	2.0	2.0000		0.0054	
b_{RESP}	0.28	0.2800		0.5980	
$ex_{ADJ,BA}$	2.0	2.0000		0.0000	
$remi_{BA}$	6.0	6.0000		0.1910	
$ex_{REFR,BA}$	0.02	0.0180		1.6500	AJ
$f_{SLCT,BA}$	0.2	0.2000		-1.3000	
$resp_{BA}^{B}$	0.01	0.0100		0.3980	
$r_{\min,BA}^{A}$	0.35	0.3500		-0.7580	
$r_{\max,BA}^{A}$	0.80	0.6404	0.611 – 0.671	0.0036	CS
$mort_{BA}$	0.02	0.0490		2.6200	AJ
μ_{PRT}	4.0	4.0000		-3.3400	
g_{PHY}	0.67	0.7011	0.600 – 0.820	0.0884	CS
g_{BA}	2.0	1.5443	1.25 – 1.91	-0.0041	CS
ex_{PRT}	0.20	0.2000		4.2900	
$f_{ex,PRT}$	0.80	0.9000		-10.8000	AJ
$resp_{PRT}^{B}$	0.01	0.0100		1.3700	
$resp_{PRT}^{A}$	0.67	0.7563	0.750 -0.762	1.7400	CS
$ex_{ADJ,PRT}$	2.0	2.0000		0.4050	
$remi_{PRT}$	4.7	4.7000		-0.5720	
pom_{PRT}	0.027	0.0270		1.4000	
μ_{MZ}	1.0	1.3000		2.6000	AJ
g_{PRT}	0.65	1.4000		0.6770	AJ
g_{TR}	1.0	0.2068	0.184 – 0.232	-0.0145	CS
ex_{MZ}	0.30	0.3000		1.4300	
$f_{ex,MZ}$	0.75	0.7500		-0.8870	
$resp_{MZ}^{B}$	0.03	0.0300		0.5810	
$resp_{MZ}^{A}$	0.25	0.2229		-0.2870	AJ
$ex_{ADJ,MZ}$	2.0	2.0000		0.0000	

$remi_{MZ}$	4.0	4.0000		-0.0058	
pom_{MZ}	0.15	0.1500		-0.0083	
$ex_{REFR,MZ}$	0.02	0.0200		-0.0369	
$remv_{MZ}$	0.8	0.8100		-0.1230	AJ
$f_{SDOM,HZ}$	0.1	0.1000		0.1670	
$f_{POM,HZ}$	0.2	0.1400		0.0853	AJ
$ex_{REFR,SDOM}$	0.0009	0.0009		-1.9700	
$q^{C}_{N,REFR}$	0.05	0.0500		-4.2000	
$q^{C}_{P,REFR}$	0.0007	0.0007		-0.0585	
$q^{C}_{N,POM}$	0.12	0.1200		3.4800	
$q^{C}_{P,POM}$	0.0045	0.0045		-1.0700	
r_{ntrf}	0.1	0.1000		0.6660	
prf_{N}	1.1	1.1000		0.5190	
prf_{P}	4.0	4.0000		-0.8920	
wnsv	2.0	4.3723	3.43 – 5.56	0.0008	CS
diss	0.02	0.0110		0.0097	AJ

Figure Legends

Figure 1. Vertically and temporally averages of modeled carbon (C), nitrogen (N) and phosphorus (P) fluxes (arrows) and stocks (boxes) normalized to C, N and P components of primary production (stocks normalized to amount of primary production during 1 day) at AS (upper panel), EQP (middle panel) and HOT (lower panel). Red arrows: microbial loop, green arrows: major food web, blue arrows: import and export including production of refractory DOM (in dashed blue arrows), yellow arrows: detritus production, dashed black arrows. Primary production is marked on top of each phytoplankton box. 1 unit C = 0.73 (AS), 0.87 (EQP) and 0.32 (HOT) mmol C m^{-3} (day^{-1}); 1 unit N = 0.10 (AS), 0.12 (EQP) and 0.045 (HOT) mmol C m^{-3} (day^{-1}); 1 unit P = 0.0096 (AS), 0.086 (EQP) and 0.0022 (HOT) mmol C m^{-3} (day^{-1}). See text for acronyms.

Figure 2. Production rate of labile DOM (LDOM) and semilabile DOM (SDOM) in carbon relative to total DOM production rate at the 3 study sites AS, EQP and HOT. LDOM is produced from bacterial mortality and excretion from phytoplankton (PHY), protozoa (PRT) and metozoa (MZ). SDOM is produced from detritus (DET) dissolution and excretion from PHY, PRT and MZ. At HOT, DOM production by PHY also includes those from unicellular N_2-fixers and *Trichodesmium*.

Figure 3. Nitrogen budgets including recycling, inputs and exports at AS, EQP and HOT. The figure shows the fluxes relative to total demand by primary production including (1) recycling to inorganic pool by metozoa (MZ), protozoa (PRT) and bacterial (BA), (2)

inputs as nitrate (NO3) from deep ocean and N_2 fixation and (3) exports as sinking detritus (DET), physical transport of semi-labile DOM (SDOM), physical transport of other biological species and production of refractory DOM. The pink diamonds indicate the imbalance of more input over export.

Figure 4. Depth-averaged semilabile DOC (SDOC) production (green line) versus consumption (black dot line) and labile DOC (LDOC) production (blue dash line) versus consumption (red dash dot line) at AS (upper panel), EQP (middle panel) and HOT (lower panel). Production in the panels is the net flux of DOC excluding bacterial consumption and therefore includes production from all the biological processes, physical transport and conversion of SDOC to refractory DOC.

Figure 5. Modeled bacterial production and bacterial growth efficiency in AS (blue), EQP (red) and HOT (black).

Figure 6. 3-day running mean anomaly of depth-integrated particulate primary production (PP) (green solid line), bacterial production (BP) (pink dash line), bacterial growth efficiency (BGE) (blue dash dot line) and the portion of bacterial carbon demand (BCD) supported by semi-labile DOM (SDOM) (black dot line) at AS (upper panel), EQP (middle panel) and HOT (lower panel). The anomalies are normalized to their mean value. The corresponding observations of depth-integrated PP (green squares) and BP (pink diamonds) are marked in the figure.

Figure 1

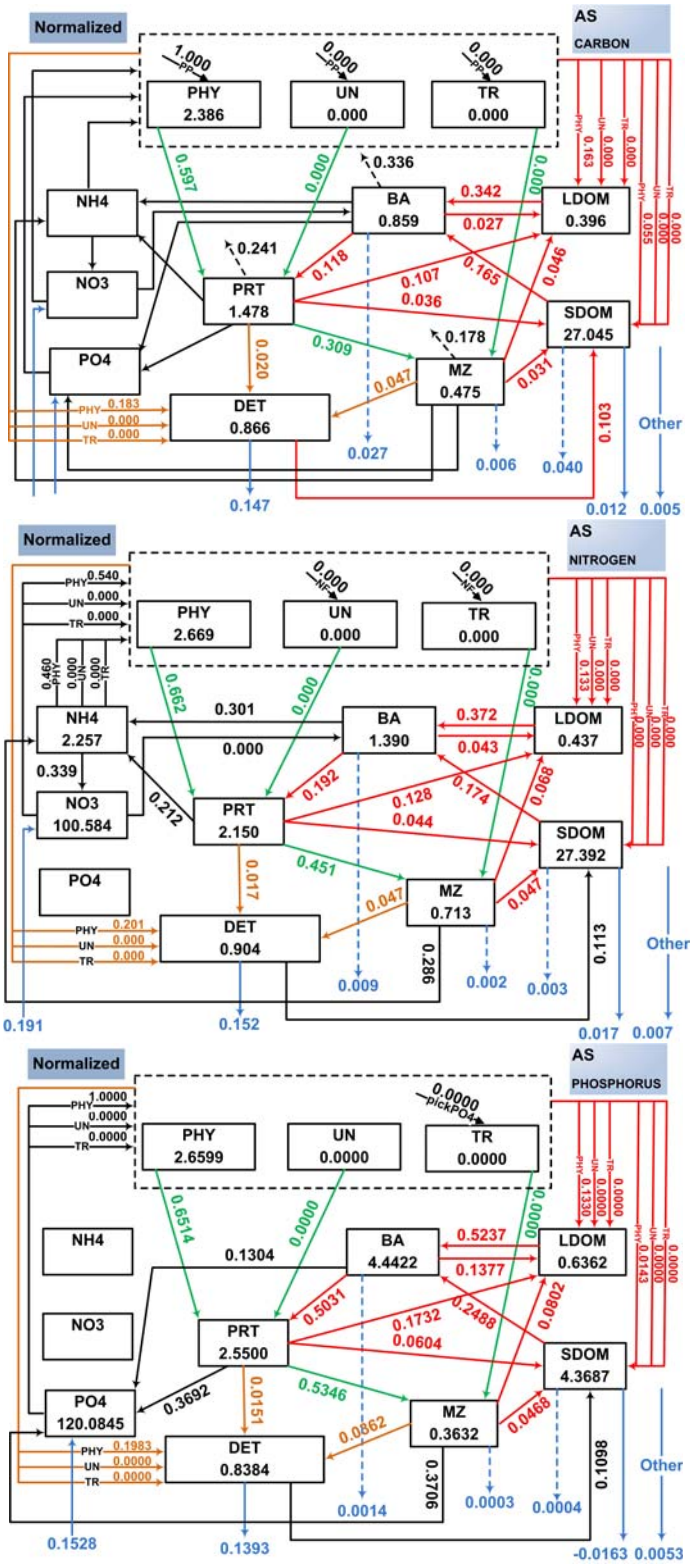

Normalized
AS
CARBON
PHY
2.386
UN
0.000
TR
0.000
NH4
NO3
PO4
BA
0.859
LDOM
0.396
PRT
1.478
SDOM
27.045
MZ
0.475
DET
0.866
Other
Normalized
AS
NITROGEN
PHY
2.669
UN
0.000
TR
0.000
NH4
2.257
NO3
100.584
PO4
BA
1.390
LDOM
0.437
PRT
2.150
SDOM
27.392
MZ
0.713
DET
0.904
Other
Normalized
AS
PHOSPHORUS
PHY
2.6599
UN
0.0000
TR
0.0000
NH4
NO3
PO4
120.0845
BA
4.4422
LDOM
0.6362
PRT
2.5500
SDOM
4.3687
MZ
0.3632
DET
0.8384
Other

Figure 1 *Continued*

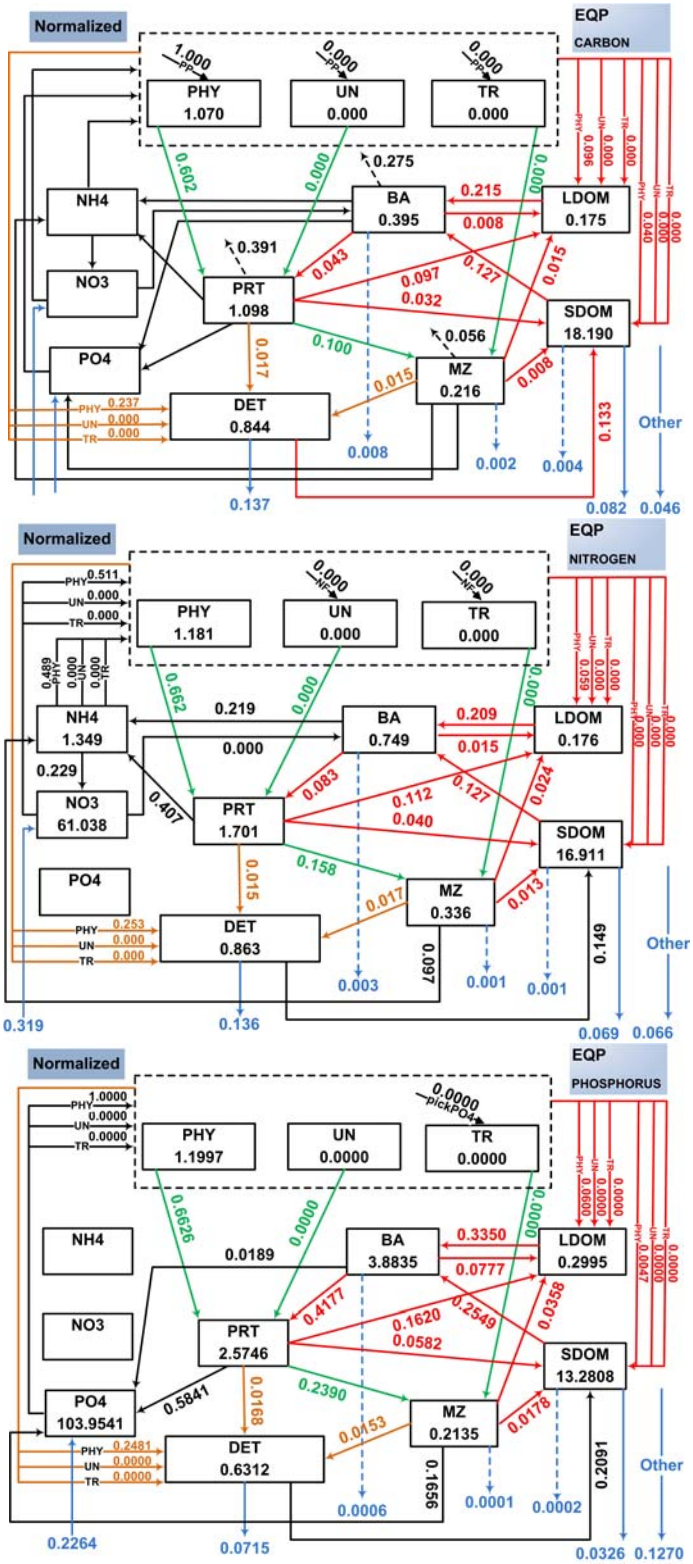


Figure 1 *Continued*

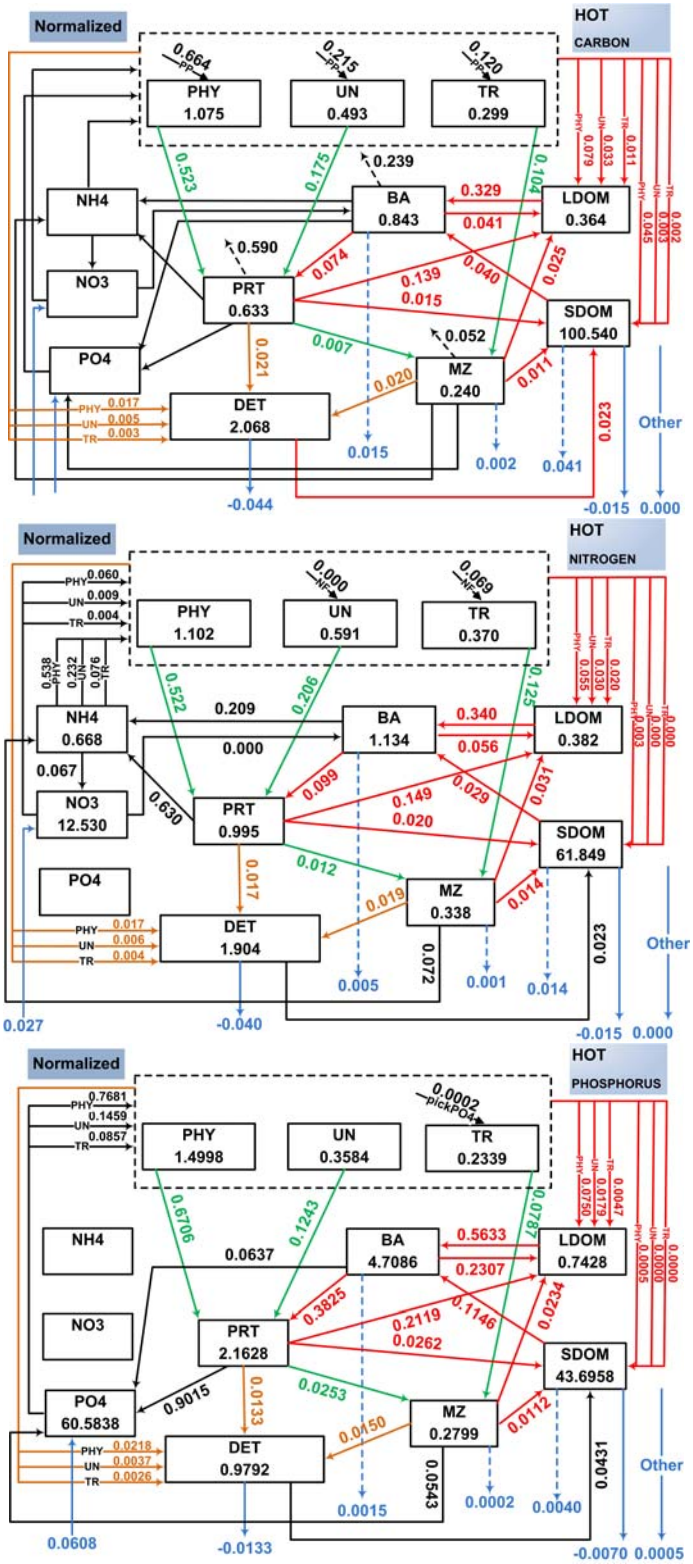


Figure 2

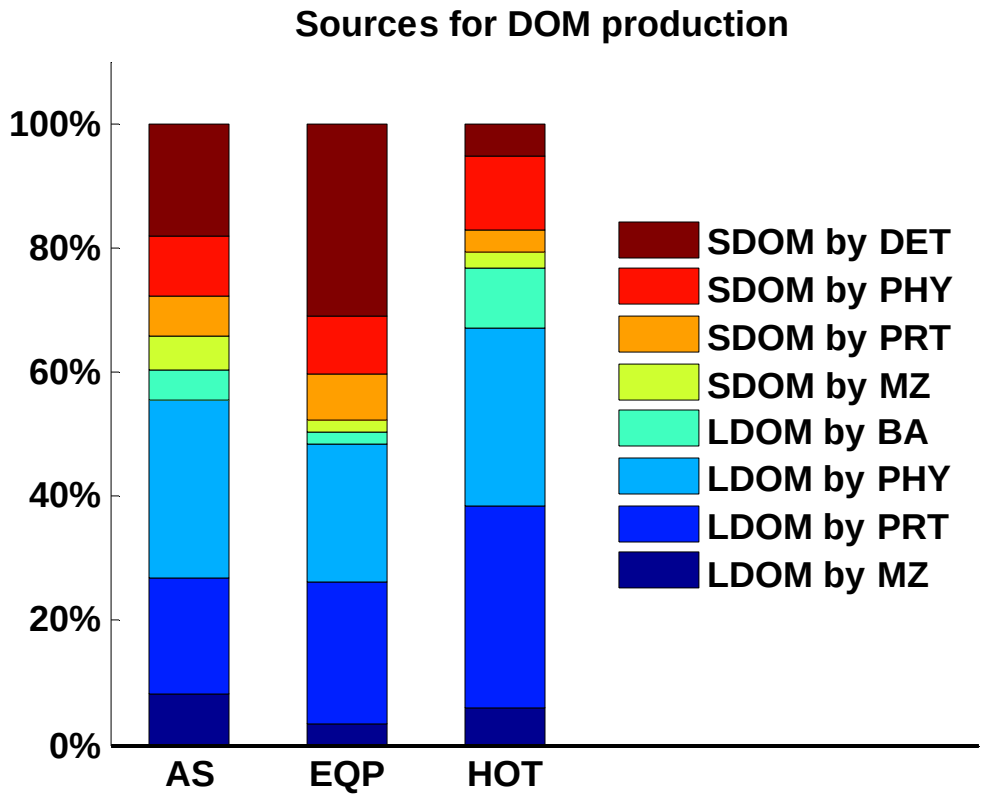

Sources for DOM production
100%
80%
60%
40%
20%
0%
AS
EQP
HOT
SDOM by DET
SDOM by PHY
SDOM by PRT
SDOM by MZ
LDOM by BA
LDOM by PHY
LDOM by PRT
LDOM by MZ

Figure 3

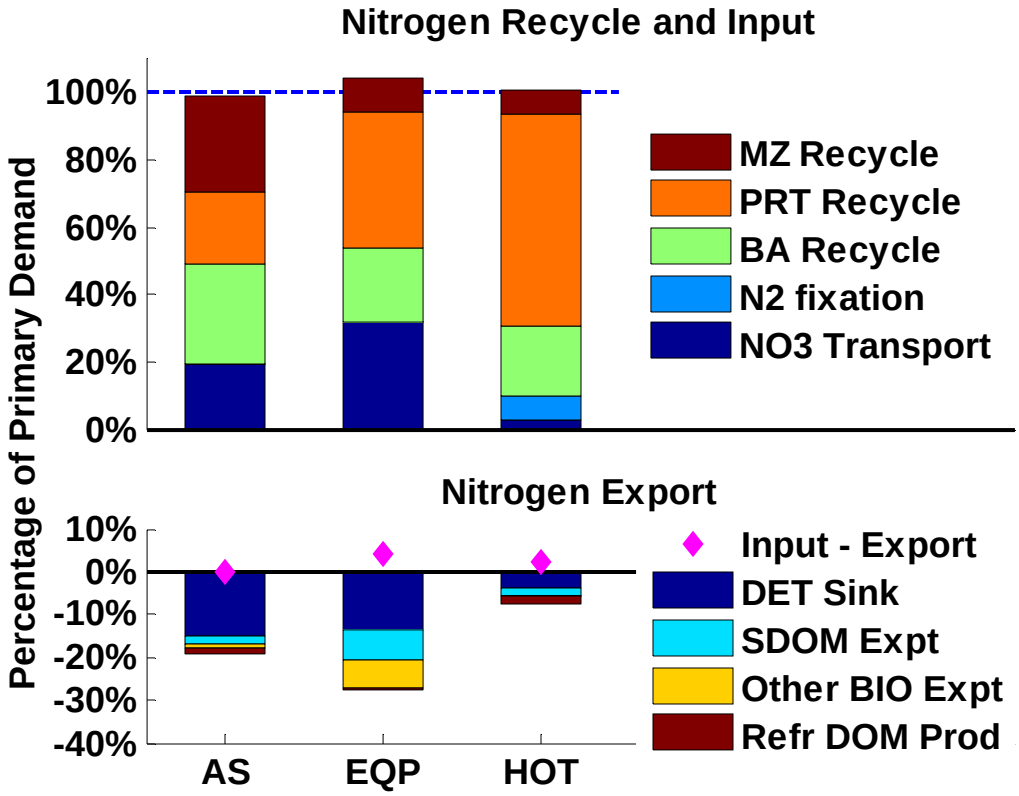

Nitrogen Recycle and Input
Percentage of Primary Demand
100%
80%
60%
40%
20%
0%
MZ Recycle
PRT Recycle
BA Recycle
N2 fixation
NO3 Transport
Nitrogen Export
10%
0%
-10%
-20%
-30%
-40%
Input - Export
DET Sink
SDOM Expt
Other BIO Expt
Refr DOM Prod
AS
EQP
HOT

Figure 4

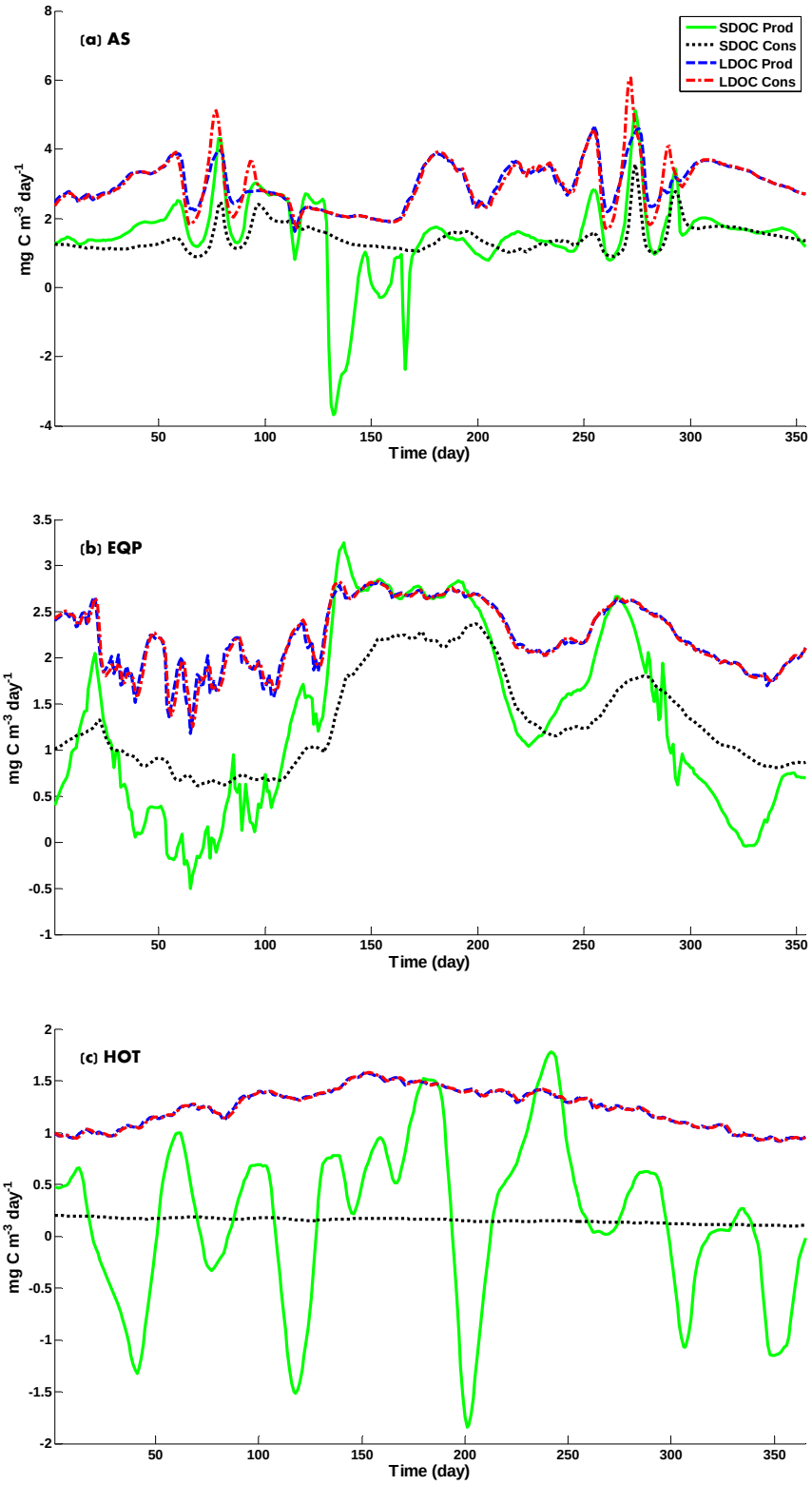

(a) AS
SDOC Prod
SDOC Cons
LDOC Prod
LDOC Cons
mg C m-3 day-1
Time (day)
(b) EQP
mg C m-3 day-1
Time (day)
(c) HOT
mg C m-3 day-1
Time (day)

Figure 5

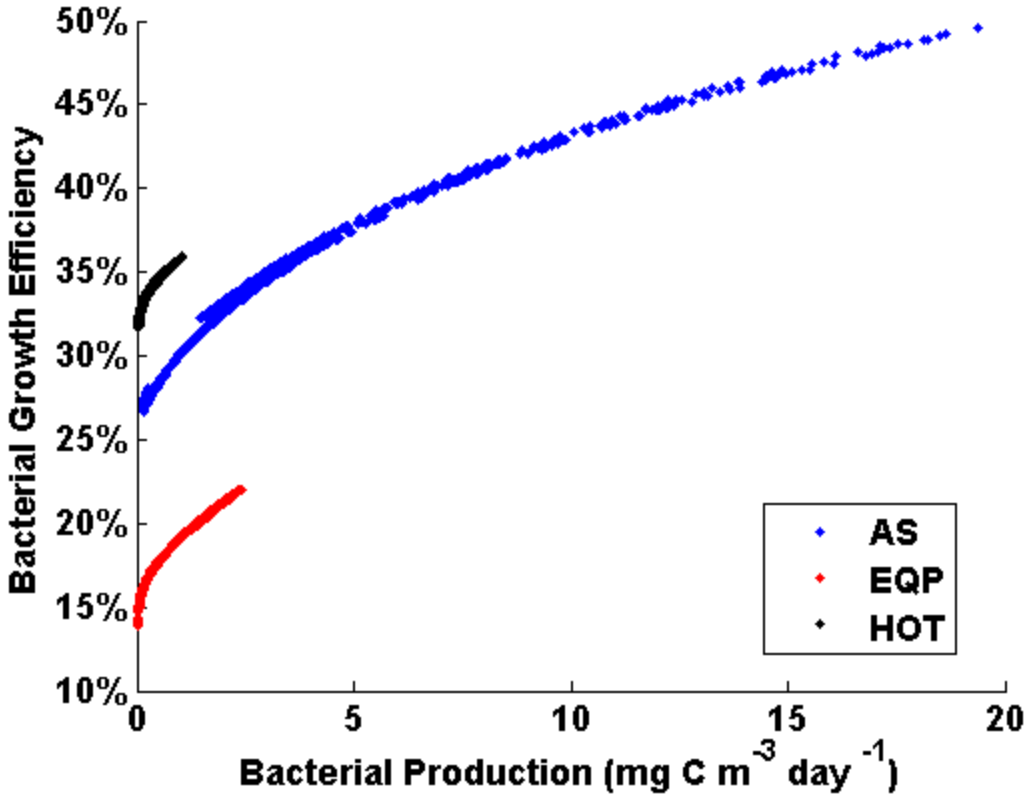
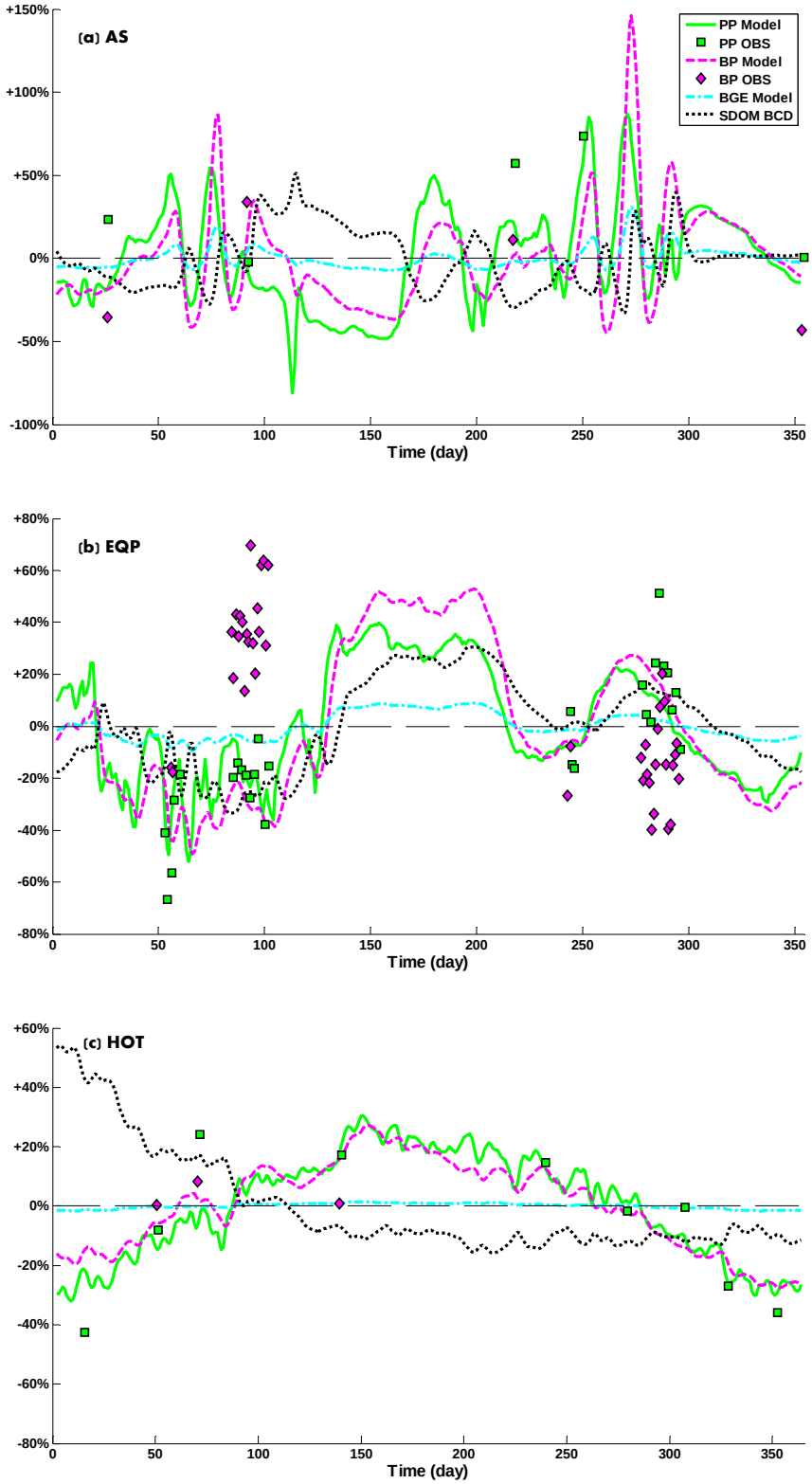

50%
45%
40%
35%
30%
25%
20%
15%
10%
Bacterial Growth Efficiency
0
5
10
15
20
Bacterial Production (mg C m-3 day-1)
AS
EQP
HOT

Figure 6


(a) AS
PP Model
PP OBS
BP Model
BP OBS
BGE Model
SDOM BCD
+150%
+100%
+50%
0%
-50%
-100%
0
50
100
150
200
250
300
350
Time (day)
(b) EQP
+80%
+60%
+40%
+20%
0%
-20%
-40%
-60%
-80%
Time (day)
(c) HOT
+60%
+40%
+20%
0%
-20%
-40%
-60%
-80%
Time (day)

Chapter 4. A 15-Year Data Assimilating Model of the Marine Microbial Ecosystem at Station ALOHA, Hawaii.

Abstract

A one-dimensional model was constructed to study the interannual variability of the marine ecosystem at Station ALOHA near Hawaii. The model was applied to the surface 200 m over the period 1991 – 2005. Data assimilation was conducted to objectively, and successfully, optimize the model to fit the 17 types of major biogeochemical observations (most with monthly profiles) during the 15 years. The model was able to simulate the interannual increase of primary production, and found different controlling mechanisms in two layers: 0 – 20 m and 50 – 70 m. The dominant mechanism for 50 – 70 m was the increased physical input of dissolved inorganic nitrogen (DIN) generated by the increasing mixed layer depth. In 0 – 20 m, together with the increased physical input of DIN, the dinitrogen (N_2) fixation was inhibited. This replacement of one new nitrogen (N) source, N_2 fixation, by another new N, physical DIN input, increased N recycling efficiency in the ecosystem and therefore enhanced the primary production. This study suggests that these two new N sources are not equal to fuel primary production. Finally, in 0 – 20 m, semilabile DON acted as a nitrogen storage pool and its production and consumption were decoupled. Part of semilabile DON usage by heterotrophic bacteria in this layer was another important form of new N, whose rate of increase equaled to one third of the rate of increase from physical DIN input.

1. INTRODUCTOIN

The long-term oceanic ecological and biogeochemical observatory, the Hawaii Ocean Time-series (HOT) Station ALOHA (22.75ºN, 158.00ºW), was established in the North Pacific Subtropical Gyre (NPSG) in 1988 (Karl et al. 2001b) as part of the Joint Global Ocean Flux Study (JGOFS) (Doney & Ducklow 2006). Since then, about 20 years of biogeochemical observations have been recorded during roughly monthly cruises. Although 20 years seems not long enough to study long-term trends or ecological shifts, Project HOT scientists have documented substantial interannual to decadal variations of the ecosystem at Station ALOHA. Corno et al. (2007) reported primary production (PP) increased ~50% in 1989 – 2004 at Station ALOHA. Although we believe that Corno et al. (2007) overestimated the rate of increase (see below), it is clear that the PP has been in an upward trend in the past two decades. Corno et al. (2007) related this PP increase to climate forcing which may have changed vertical stratification and nutrient supply and may also have altered the composition of the plankton assemblage. The datasets from Station ALOHA also potentially indicate a development of dinitrogen (N_2) fixation, revealed by the observed increasing nitrogen (N) to phosphorus (P) ratios of dissolved organic matter, suspended particles and exported particles over the years (Karl et al. 1997). The N_2 fixation has been estimated to support a significant amount of new PP at Station ALOHA (Karl et al. 1997, Casciotti et al. 2008) and may enhance carbon sequestration (Karl & Letelier 2008). Recent 3-year direct measurements of N_2 fixation rates and community at Station ALOHA revealed that mesoscale physical forcing may modulate the variability in N_2 fixation (Church et al. 2009).

N_2-fixing species (diazotrophs) can contribute to PP both directly by their own

growth and indirectly through recycling the newly fixed N to bioavailable forms such as ammonium and urea. Not all the newly fixed N will stay in the surface ocean because of vertical export, and so it is necessary to evaluate the efficiency of the recycling of the newly fixed N. Newly fixed N can recycle through multiple pathways including the extracullular release of ammonium (Mulholland 2007) and dissolved organic nitrogen (DON) (Capone et al. 1994, Glibert & Bronk 1994), cell lysis (Hewson et al. 2004) and grazing by zooplankton (O'Neil et al. 1996). The released DON will be mainly consumed and recycled by heterotrophic *Bacteria* and *Archaea* (termed heterotrophic bacteria hereafter) before the N becomes bioavailable to other trophic levels (Carlson 2002). In addition, although zooplankton can release ammonium directly, they also excrete DON, which in turn will be processed by heterotrophic bacteria. Heterotrophic bacteria have their own dynamics affecting the efficiency and speed that DON is recycled. Thus to study the relationship between development of N_2 fixation and supposed enhanced PP, heterotrophic bacteria are one of the key link points.

Besides their role for linking N_2 fixation and PP, heterotrophic bacteria are major players in ocean carbon and nutrient cycles. About 50% of PP by phytoplankton in oceanic ecosystems is released as dissolved organic matter (DOM) (Ducklow & Carlson 1992, Williams 2000), and DOM is one of the largest exchangeable carbon reservoirs on earth (Hedges 1992). Thus as the major consumer of the DOM pool, heterotrophic bacteria are a key component in the global carbon cycle. Together with zooplankton, heterotrophic bacteria are also the major sources for recycling organic N and P to inorganic forms in the ocean (Kirchman 2000) and indirectly support PP. At Station ALOHA, observations have revealed multiyear increases of DOM in 1989 – 1999

(Church & Ducklow 2002) and temporal and vertical dynamics of heterotrophic bacterial production (Church et al. 2006) in 2000 – 2005. But it is not clear what processes controlled these dynamics.

To address these issues, we constructed a 1-dimensional marine ecosystem model (Chapter 2) to apply to Station ALOHA over 1991 – 2005. With roughly monthly observations available from HOT program, we conducted data assimilation to objectively optimize model parameters in order to minimize the differences between 17 types of major biogeochemical observations and their model equivalents. We worked to simulate the interannual variability of PP, N_2 fixation and heterotrophic bacterial dynamics at Station ALOHA, and investigated the mechanisms controlling these decadal changes.

2. METHODS

The ecosystem model simulates carbon (C), N and P cycles independently through the state variables including normal phytoplankton (PHY), N_2-fixing *Trichodesmium* spp. (TR) and unicellular diazotrophs (UN), heterotrophic bacteria (BA), protozoan (PRT) and metazoan (MZ) zooplankton, labile DOM (LDOM) and semilabile DOM (SDOM), detritus (DET) and inorganic nutrients (ammonium NH4, nitrate NO3 and phosphate PO4). Chlorophyll *a* of PHY, TR and UN is also simulated separately from the biomass. Thus the ratios among C, N, P and Chlorophyll *a* are not constant but vary over time and space. The full model structure is described in Section 2 of Chapter 2.

The ecosystem model was then applied to a 1-dimensional framework for Station ALOHA (22° 45′ N, 158° W) for 1991 – 2005. The length of each time step is 1 hour. A spin-up period of 455 days (5 seasons) is run ahead of the simulation to minimize the

influence of initial conditions. A total of 25 vertical layers were used, with 10 surface layers of 5 m each and 15 bottom layers of 10 m each, the model domain covering the upper 200 m of the water column. The model is forced by sea surface photosynthetically active radiation (PAR), mixed layer depth, water temperature, vertical velocity, vertical eddy diffusivity, as well as bottom boundary conditions including physical exchanges of nitrate, phosphate and semilabile DOM. The full description of the model setup is in Section 3.1 of Chapter 2.

Data assimilation was conducted to minimize the misfits between observations and modeled results by objectively optimizing the model parameter values. A variational adjoint scheme (Lawson et al. 1995, Friedrichs et al. 2007) was used for this process. The purpose of the scheme is to optimize parameters of model equations in order to minimize the cost function J, which is a metric to evaluate the difference between modeled results a and observations $\hat{a}$. J summarizes the average Chi-squared value (χ^2) of each data type:

$$J = \sum_{m=1}^{M} \frac{1}{N_m} \sum_{n=1}^{N_m} \left(\frac{a_{m,n} - \hat{a}_{m,n}}{\sigma_m} \right)^2 = \sum_{m=1}^{M} \overline{\chi_m^2}, \quad (1)$$

where m and n represent assimilative data type and data points, M and N are the total numbers of assimilative data types and data points; σ is the target error whose value for each data type can be found in Table 1 of Chapter 3. Thus a lower J value indicates a better fit of the model to observations. Not all the parameters are constrainable by the assimilated data. Only a small set of parameters was optimized based on diagonal elements of the inverse Hessian matrix (Friedrichs et al. 2007). The observations assimilated for Station ALOHA include nitrate, phosphate, metazoan zooplankton C

biomass, phytoplankton N biomass and Chlorophyll *a*, PP rate, heterotrophic bacterial C biomass and production, semilabile dissolved organic C, N and P, particulate organic C, N and P and particulate C, N and P collected at 150 m sediment traps. More details of data assimilation are described in Section 3.2, Chapter 2.

3. Results

3.1. Model Optimization

The data assimilation was able to optimize 27 parameters with different uncertainties. I evaluate a parameter as well constrained if its uncertainty (one standard deviation) is below 40% of its optimized value. Thus 12 parameters were well-constrained, and the other 15 parameters were optimized but not well-constrained by data assimilation **(Table 1)**. During the optimization, reductions occurred for the contributions of most data types to the overall cost function J and the value of J was minimized **(Figure 1)**. However, because the target of the data assimilation was to minimize the sum of all the costs instead of each individual component, some components were not reduced and were even increased, although the total cost was minimized. The cost of nitrate (NO3) had a small increase, and the cost of particulate organic phosphorus (POP) was about doubled. Most other costs were dramatically reduced.

The well-constrained parameters included those controlling (1) photosynthesis, (2) predation by protozoan and metazoan, (3) DOM lability, (4) respiration of heterotrophic bacteria and protozoa and (5) production and sinking velocity of particulate organic matter (POM), as the assimilated data types had direct relations to these parameters. (1)

Photosynthesis related to primary production, phytoplankton biomass and chlorophyall *a*, (2) predation related to phytoplankton and metazoan zooplankton biomass, (3) DOM lability related to DOM concentration and heterotrophic bacterial production, (4) respiration of heterotrophic bacteria related to heterotrophic bacterial biomass and production, (5) POM production and sinking related to POM concentration and flux. Protozoan respiration might relate to multiple factors as protozoa were in the center of the food web. Biomass of phytoplankton, heterotrophic bacteria and metazoa and nutrient levels (protozoa were a major remineralizer) can potentially constrain protozoan respiration. Although many other parameters were not optimizable, the gradients of the cost function J with respect to these parameters were mostly low **(Table 1)**. The gradients had a special definition in this study in order to address the response of the cost function to the relative change of parameters:

$$g = \frac{J\left(x \cdot e^{\pm a}\right) - J\left(x\right)}{\pm a}, \qquad (a \text{ infinitely small}). \tag{2}$$

As $e^{\pm a} \approx 1 \pm a$ when a is small enough, for instance, a 10% change of optimized parameter values will only result in a cost function change equaling 10% of gradient. Most gradients were less than 1 and all of them were less than 10 **(Table 1)**, thus a 10% change of an optimized parameter value would only result in the cost function change of less than 1 (mostly less than 0.1). A change of cost function is not significant if it is less than the number of observation types (17 here) (Friedrichs et al. 2007). Therefore the low gradients indicated J approaches at least a local minimum, even though the global minimum cannot be guaranteed. However, the reasonable values of all the parameters consistent to literature ranges and the reduced differences between the model results and

observations revealed that the data assimilation successfully improved the model results.

3.2. Fit of Specific State Variables and Rates

The depth-integrated observations and their model equivalents were compared in **Figure 2**. Because only the model equivalents, i.e., the subsampled results at same depths and times as observations, were integrated over depth, **Figure 2** gave a direct evaluation of the model skill to a subset of the assimilated observations. The assimilated observations of nitrate, phosphate and metazoan zooplankton C biomass were their average concentrations in the mixed layer (Section 3.2, Chapter 2), thus the comparisons of these variables in **Figure 2** were also against their average values in the mixed layer. The sinking particulate C, N and P fluxes were compared at 150 m depth in **Figure 2**. The depth-integrated results showed that many variables had good skill to the corresponding observations. Most variables had correct seasonal variation patterns, even if there is a bias in the mean relative to the observations.

The modeled nitrate level in the mixed layer stayed below 0.01 μM, consistent with most observations, but did not catch several observed spikes. The observed phosphate level in the mixed layer showed a significant decrease in 2001. However, the modeled phosphate in the mixed layer was much lower than the observations before 2000. It suddenly increased in 2000 to the observed level. In 2001-2002 it dropped in the same period when the observations dropped, and reached the observed low levels in 2003-2005.

The modeled metazoan zooplankton biomass in the mixed layer was quite comparable to the observations. The depth-integrated modeled phytoplankton N biomass had much lower variations, both seasonally and interannually, than the observations. But

its average level over the whole period was comparable to the observed average. The phytoplankton N biomass observations were derived from the cell number counts with constant conversion factors over the whole period for *Prochlorococcus*, *Synechococcus* and Eukaryotes respectively (see details in Chapter 2). The actual conversion factor values change if the community changes. Thus the observed significant interannual variations of phytoplankton N biomass could sometimes be unrealistic, because they came from indirect measurements. This was partially supported by the better fit of the depth-integrated Chlorophyll *a*, which was directly measured. The depth-integrated PP showed a very good fit both seasonally and interannually, except for 2002 when the modeled results were lower than observations. The depth-integrated heterotrophic bacterial biomass underestimated the observed magnitude of the seasonal and interannual variations, although the average level was reasonable. Like phytoplankton biomass, the constant conversion factor from heterotrophic bacterial cell counts to their biomass may contribute to the observed variations.

Heterotrophic bacterial production (BP) was rarely measured (totally 13 profiles) so that the seasonal variations cannot be revealed from the few observations. There was a 2-year gap in 2002-2004 when no BP was measured, and the depth-integrated BP observations were significantly different before and after the gap. Because a constant conversion factor from the leucine incorporation rates to BP was used for all the profiles, we do not know if this difference reflected the real conditions, or it resulted from uncertainties of the conversion factor. The modeled depth-integrated BP was lower than the observations before the gap and was higher after the gap. But the modeled average was comparable to the observed average.

The depth-integrated modeled semilabile dissolved organic carbon (DOC) was in the range of the observations, but it did not successfully simulate interannual variations. There was evidence that the semilabile DOC measurements were overestimated before 2001 (Matthew Church at University of Hawaii, personal communication); my correction for this error (see details in Chapter 2) may have resulted in lower semilabile DOC observations after 2001 than the years before. Thus although the modeled semilabile DOC was in the range of the previous years, it showed an overestimate when compared to the lower observations after 2001. The depth-integrated modeled semilabile dissolved organic nitrogen (DON) had a good fit to the observations, except for 1994 when the model underestimated them. The depth-integrated modeled semilabile dissolved organic phosphorus (DOP) varied less than the observations, but it was within the observed range.

The depth-integrated modeled particulate organic carbon (POC) had correct seasonal variations compared to the observations. The modeled POC was within the observed range, but the modeled average was below the observed average. The depth-integrated modeled particulate organic nitrogen (PON) had good skill relative to the observations, especially after 1997, but the model cannot simulate the strong seasonal variations before 1997. The depth-integrated modeled particulate organic phosphorus (POP) was stable over the entire model period, while the observed POP showed a steady downtrend over the 15-year period with a ~ 40% decrease. However, there was no similar pattern in particulate phosphorus flux collected by sediment traps.

3.3. Average Modeled Results in the 15-year Period

Figure 3 shows the 15-year-average C, N and P stocks (in boxes) and fluxes (arrows)

after data assimilation. The fluxes in **Figure 3** were normalized to C, N and P components of PP respectively, and the stocks were normalized to 1-day of PP. The PP used here includes both particulate and dissolved PP. The average results were analyzed in Chapter 3 for three oceanic sites over a single year of observations, including Year 2002 at Station ALOHA. Thus here I only discuss briefly some key findings of the average results.

In the model, normal phytoplankton contributed 84% of total carbon-based PP, and N_2-fixing diazotrophic phytoplankton contributed 16%, including 10% from *Trichodesmium* spp. and 6% from unicellular diazotrophs (**Figure 3a**). About 90% of PP was respired by bacteria and zooplankton. The remaining 10% was exported in form of sinking detritus (5.4%), semilabile DOC transport (3.5%) and production of refractory DOM (1.7%). Heterotrophic bacterial carbon demand (BCD) was 36% of PP. 86% of BCD was supported by labile DOC, and 14% was supported by semilabile DOC. Heterotrophic bacteria respired 63% of BCD, making bacterial growth efficiency (BGE) 37%. The total respiration and export balanced to PP, which indicated the total C inventory did not change significantly in the 15 years.

Physical transport of nitrate and N_2 fixation supported 2.3% and 6.1% of the total N requirement by PP respectively (**Figure 3b**). That is, the f-ratio, the fraction of the PP supported by new N, was 8.4%. The model result showed that N_2 fixation dominated the supply of new N, as discussed below. N export equaled the f-ratio (8.4% of PP), including sinking detritus (5.2%), semilabile DOC transport (2.5%) and production of refractory DOM (0.7%). The majority of the N requirement for PP was supplied by N recycling from zooplankton (72% of PP requirement) and by heterotrophic bacteria (19%

of PP requirement). The total N input and export were about in balance with each other, i.e., the total N inventory did not change significantly in the 15 years.

In the simulation, 2.1% and 0.2% of the P requirement by PP was supported from physical phosphate transport and the active P acquisition by *Trichodesmium* spp., respectively (**Figure 3c**). Along with the low new P supply, the P export rate was also low, which indicated that P recycled more efficiently than N. It was realized mainly by the extremely low P export through sinking detritus. Interestingly, the model predicted that more P was exported in physical exchange of semilabile DOP than in sinking detritus. P recycled by heterotrophic bacteria only supported 12% of the PP requirement, while N they recycled supported 19% of PP requirement. It resulted from the higher P:N ratio of heterotrophic bacteria (0.11) than zooplankton (0.04). Thus if heterotrophic bacteria and zooplankton recycle same amount of N, heterotrophic bacteria will store more P in their cells and recycle less. The total P input and export were about in balance with each other, i.e., the total P inventory did not change significantly in the 15 years.

3.4. Interannual Variation of Primary Production

The observed PP showed an overall significant uptrend from 1991 to 2005 (linear regression, $p = 0.047$, **Figure 4**). The Model-I linear regression predicts the increase rate of PP at 4.77 mg C m^{-2} day^{-1} per year, with 27% of the observed interannual variability of PP can be explained by time. The modeled PP also showed a significant increase in 1991 – 2005 (linear regression, $p = 0.002$, **Figure 4**), at a Model-I linear-regression predicted rate of 2.45 mg C m^{-2} day^{-1} per year. 55% of the observed interannual variability of PP can be explained by time. Although the linear-regression predicts different increase rates for the observation and the model, the absolute PP gain over 15 years were about same

for the observation and the model. The observed PP had increased from 498 to 540 mg C m^{-2} day^{-1}, i.e., an 8.4% increase; and the modeled PP had increased from 477 to 525 mg C m^{-2} day^{-1}, i.e., a 10% increase. The interannual variation of the modeled PP was weaker than that of the observations. But the modeled PP was close to the observations in most years, except for 2002 when the model cannot reach the extraordinary high PP observations. Comparing the full profiles of PP (**Figure 5**), there was no extremely high PP rate observed in 2002. But in 2002, PP started to increase earlier and declined later than other years. In addition, the euphotic zone in 2002 was deeper than other years. The modeled PP cannot reproduce this pattern for 2002 (**Figure 5**). The observed nitrate level in the mixed layer around the beginning of 2002 was extremely high (**Figure 2**) due to a winter mixing event. The model could not simulate this event, probably because of the limitation of its physical scheme lacking mesoscale and sub-mesoscale features. Once this new nitrate was imported into the mixed layer and assimilated into phytoplankton, active recycling maintained the new N in the surface to support the primary production over the whole year of 2002.

The interannual increase of the modeled PP was mostly contributed by normal phytoplankton with a 40 mg C m^{-2} day^{-1} increase in the 15 years. A small increase of 7 mg C m^{-2} day^{-1} was contributed by unicellular diazotrophs (**Figure 6a**). The modeled PP of *Trichodesmium* spp. did not show strong interannual variability and only contributed 1 mg C m^{-2} day^{-1} to the total PP increase. The modeled PP of unicellular diazotrophs had highest relative increase of 34% in the 15 years; normal phytoplankton increased 10% and *Trichodesmium* spp. only increases 2% (**Figure 6b**). The increase of the modeled PP occured mostly in the surface 90 m for both normal phytoplankton and unicellular

diazotrophs, as indicated by the linear regression for each vertical layer (**Figure 7**). The linear regression for each month revealed that the modeled PP had its highest increase rate in June, July and August for normal phytoplankton (**Figure 8a**), and in April, July, August and September for unicellular diazotrophs (**Figure 8b**).

In the following, I analyze the long-term PP increase in two separate layers: (1) 0 – 20m, and (2) 50 – 70 m. The surface 20 m was always within the mixed layer, and the model mixed all the materials in the mixed layer at every time step. Therefore the analysis at the surface 20 m can represent the condition for the whole mixed layer. It also represents the nearly uniform increasing rate of PP in the surface 50 m (**Figure 7**). The PP had the highest rate of increase in the 50 – 70 m layer (**Figure 7**). The mixed layer depth at Station ALOHA had average depth ~ 60 m, and usually reached 20 – 40 m in the summer and 80 – 100 m in the winter. Thus the 50 – 70 m layer can be seasonally and/or interannually below or above the mixed layer depth. The physics, and therefore the mixed layer depth and nutrient levels, could heavily influence the PP in 50 – 70m. As shown in **Figure 6**, simulated *Trichodesmium* spp. did not increase significantly over time. Thus my analysis below will mostly focus on normal phytoplankton and unicellular diazotrophs.

3.4.1. Layer of 0 – 20 m

In 0 – 20 m, the PP (C component) of both normal phytoplankton and unicellular diazotrophs increased significantly over years ($p < 0.01$ and $p = 0.02$ respectively) (**Figure 9 a, b**). PP can be limited either by photosynthetically active radiation (PAR) or by ambient nutrient concentrations. For the macro nutrients N and P, PP at Station ALOHA was limited by N in the surface as revealed by low dissolved inorganic N:P ratio

(Karl et al. 2001a), which was also confirmed by the model. Thus the interannual increase of PP at Station ALOHA could be contributed either by the increase of PAR, or by the increase of inorganic dissolved N, including nitrate and ammonium. If PAR was the major factor, the modeled C/N of PP should show a significant interannual increase, as the model allowed flexible stoichiometry. The model simulated C and N assimilation quasi-independently. C assimilation rate depended on PAR and N assimilation depended on N supply. Unless cellular C/N ratio approached the maximum or minimum allowable value, the assimilation rates were not intervened (see details in Chapter 2). Thus if PAR was the limiting factor and its increase was the reason for increasing PP, the model would assimilate more C relative to N. That is, the same amount of N supply could support more C assimilation. However, the modeled C/N ratio of PP did not show a significant interannual increase for both normal phytoplankton ($p = 0.65$) and unicellular diazotrophs ($p = 0.98$) (**Figure 9 c, d**). Also, the annual averages of the PAR used to force the model did not show significant increase over 1991 – 2005 (**Figure 10**), and the level in 2005 was even lower than that in 1991. As expected, the N component of PP in 0 – 20 m significantly increased over the years for both normal phytoplankton ($p = 0.03$) and unicellular diazotrophs ($p=0.05$) (**Figure 9 e, f**). Thus an increase in the inorganic N supply supported the interannually increasing PP.

I first compare the differences of the annual average stocks and fluxes in 0 – 20 m between 1991 and 2005, i.e., the net change of the ecosystem status in the whole period. The increased N component of PP in 0 – 20 m was supported by the increased nitrate and ammonium concentration (**Figure 11**). Although the higher nitrate and ammonium concentrations are necessary for instantaneous increases in PP, sustained increases require

higher supply rates (both biologically and physically) of nitrate and ammonium. Over the long term and at steady-state, the production and consumption of biological species or chemical substrates are in balance. Any small imbalance will result in a significant reduction or accumulation over a 15-year simulation. For instance, the relatively small accumulation of nitrate and ammonium in the 15 years (**Figure 11**) indicates that the balance for their production and consumption was strong.

In **Figure 12**, I compare the difference for two years (2005 minus 1991) in the balances of all the production and consumption rates of nitrate and ammonium in 0 – 20 m. As expected, the consumption of nitrate and ammonium by the three primary producers increased in 2005 relative to 1991. The physical transport of nitrate and ammonium into the layer increased by 1.09 μmol N m^{-3} day^{-1}. However, another form of new N supply, the N_2 fixation by *Trichodesmium* spp. and unicellular diazotrophs, declined by 1.24. Thus the net supply of new N declined by 0.15. The interannual variations of physical transport of dissolved inorganic N (DIN, including nitrate and ammonium) and N_2 fixation are shown in **Figure 13**, which shows how the physical DIN input inhibits N_2 fixation in the model. Meanwhile, the total N consumption by primary producers increased. Thus if there was no other new N imported, the modeled interannual increase of PP was associated with a lower f-ratio and more efficient N recycling (discussed later), not a net increase in new N supply.

Although physical import of nitrate and biological import of fixed N_2 are both considered as new N input, it seems that a higher N recycling efficiency is preferred when physical nitrate import dominates, and a lower efficiency is related to N_2 fixation. In **Figure 14**, I estimate the N recycling efficiencies when N is consumed by the different

primary producers. The figure represents the modeled average status of 0 – 20 m in 1991 - 2005. After inorganic N is consumed in forms of nitrate, ammonium and N_2 fixation, it can be exported (1) directly through the mixing of primary producers, indirectly (2) through detritus and (3) through semilabile DOM. The unexported N will be remineralized back to ammonium through protozoan and metazoan zooplankton or through the DOM – heterotrophic bacteria pathway. In one cycle, that is, starting from N consumption by primary producers and ending when N returns back to the ammonium pool, 8.4%, 5.7% and 21.5% of N will be exported through the pathways of normal phytoplankton, unicellular diazotrophs and *Trichodesmium* spp., respectively. Unicellular diazotrophs have the highest N recycling efficiency, slightly higher than normal phytoplankton. That is mainly because a larger amount of N of unicellular diazotrophs is excreted as labile DOM (data not shown), a more efficient pathway, and thus a smaller amount is grazed, compared to normal phytoplankton. *Trichodesmium* spp. has much lower N recycling efficiency than both unicellular diazotrophs and normal phytoplankton. One main reason is that *Trichodesmium* spp. is consumed by metazoan zooplankton, which have low recycling efficiency and transfer a large amount of N to detritus and semilabile DOM. Unicellular diazotrophs and normal phytoplankton are consumed by protozoan zooplankton, which have much higher recycling efficiency. Another reason is the high direct export of *Trichodesmium* spp. through mixing. The growth of *Trichodesmium* spp. occurs mainly in the surface because of its high light-dependence. Thus the physical scheme of the model mixes it in the mixed layer, and results in a large amount of export out of the 0 – 20 m layer. As the model does not consider the self-buoyancy of *Trichodesmium* spp., this export could be an artifact. However, even if

only the export from detritus and semilabile DOM is considered, the N recycling efficiency of *Trichodesmium* spp. is still much lower than the other two phytoplankton species. Although unicellular diazotrophs have even higher N recycling efficiency than normal phytoplankton, their N_2 fixation only accounted for ~ 14% of total N_2 fixation in the modeled results, and the other 86% was produced by *Trichodesmium* spp. with low N recycling efficiency. In summary, N recycling efficiency can be enhanced if N_2 fixation is replaced by physical import of nitrate, which could possibly result in higher primary production.

The modeled ammonium production rate in 0 – 20 m was higher in 2005 than in 1991 (**Figure 11**), consistent with the expectation of enhanced N recycling. The increase was mostly contributed by protozoan zooplankton with higher grazing rates on phytoplankton (**Figure 12, 15**). The heterotrophic bacteria also contributed an increase of 0.86 μmol N m^{-3} day^{-1}, resulting mainly from the increased production rate of labile and semilabile DON (**Figure 12, 16**). Protozoan and metazoan zooplankton contributed an increase of 0.86 to labile DON, and phytoplankton contributed an increase of 0.22 to labile DON (**Figure 16**). But the ingestion of heterotrophic bacteria by protozoan zooplankton also increased by 0.76, thus the net contribution by zooplankton was 0.10 and the total contribution from labile DON pathway was 0.31 (**Figure 16**). The other pathway, semilabile DON, contributed the increase by 0.49 (**Figure 16**). However, this increase was not balanced with the change of the biological production of semilabile DON. The biological production of semilabile DOM from phytoplankton, zooplankton and detritus together increased by 0.12 (**Figure 16**), which was lower than the 0.49 increase of N consumption from semilabile DON. This unbalance was not generated by

the increase of physical import of semilabile DON. Actually the physical EXPORT of semilabile DON increased by 1.38 (note the direction of the arrow) (**Figure 16**). In the above discussions related to **Figure 12, 15 and 16**, the residence time for most substrate and species was short enough that I assumed the production and loss was tightly balanced, thus an increased production meant an increased consumption (loss). In contrast, the modeled residence time for semilabile DOM at Station ALOHA was 3 to 4 years (Figure 3, also Table 4 in Chapter 3), and the production and consumption (loss) were not necessarily in tight balance over a time scale of one year. The simple comparison between two years, 1991 and 2005, was not adequate to provide useful information as to why the N flux from semilabile DON to heterotrophic bacteria can increase.

The production of semilabile DON, including the biological production and physical exchange, had high interannual variability (**Figure 17 a**), which was mostly from the physical exchange (**Figure 17 b**). The consumption of semilabile DON by heterotrophic bacteria was relatively stable interannually (**Figure 17 a**). Thus the net flux, the difference between the production and consumption of semilabile DON, can be positive or negative (**Figure 17 c**). In 1991 – 1994, the positive flux continuously increased the stock of semilabile DON; the strong deep mixing in 1995 and 1999 set up a negative net flux and the stock became lower; the constant stratification in 2000 – 2003 set up a strong positive net flux and the stock increased; and the deep mixing and negative net flux in 2004 – 2005 set the stock lower (**Figure 17 d**). But the variation of semilabile DON only accounted for up to 15% of the total stock, and the consumption by heterotrophic bacteria was not significantly influenced by, or correlated to, the stock and the flux of semilabile DON.

When semilabile DON is produced by biological species and then consumed by heterotrophic bacteria at the same rate of the production, the flux is a recycled N. But if the consumption rate of semilabile DON is higher than its biological production rate, in a short time scale this difference can be considered as another form of new N. This form of new N was 0.49 – 0.12 = 0.37 μmol m^{-3} day^{-1} greater in 2005 than in 1991 (**Figure 16**), which equals to 34% of the increase of physical DIN input (**Figure 12**). Along with the 1.09 increase of physical DIN input and 1.24 reduction of N_2 fixation (**Figure 12**), the total new N supply increased by 0.22 in 2005 over the level in 1991. The new N supply would have declined by 0.15 if the supply from semilabile DON was not considered.

The concept of new N, the N that supports new production, is conventionally defined as external sources of N imported into the euphotic zone (Dugdale & Goering 1967). Current paradigm includes following new N sources: nitrate diffusion and convention from deep water, N_2 fixation, river runoff and atmospheric deposition (Mulholland & Lomas 2008). This paradigm is rational in long time scales such as more than decades. However, this study identified another new N source, semilabile DON, on shorter time scale. Especially in an oligotrophic system where the residence time of semilabile DON is in order of several years, longer than most biological time scales, the production and consumption of semilabile DON is not coupled. Thus the net consumption of DON can support new primary production on annual to decadal time scale. If this source is ignored, the N budget on this time scale could be unbalanced: the total N import from the conventional sources described above would be less than N export. The new N supplied by semilabile DON can be considered as a storage pool: the organic N produced in previous years can be stored in the semilabile DON pool and be used as a new N

supply in the future.

In summary, the modeled interannual PP increased in 0 – 20 m at Station ALOHA was mainly generated by (1) the replacement of partial N_2 fixation with physical nitrate import, which enhanced the N recycling efficiency, and (2) the release of previously accumulated N from the semilabile DON pool.

3.4.2. Layer of 50 – 70 m

The C/N component of PP in 50 – 70 m did not increase and even slightly decreased over time (**Figure 18**). Similar to 0 – 20 m, this indicated that N supply, not the PAR, was the major factor determining the interannual PP increase in 50 – 70 m.

Similar to 0 – 20 m, the physical input of nitrate and ammonium to the 50 – 70 m layer increased significantly, which inhibited the N_2 fixation rates (**Figure 19**). The replacement of part of N_2 fixation by physical input of new N could enhance the N recycling efficiency, as discussed above. However, the increase of physical import of new N in 50 – 70 m was much larger than the reduction of N_2 fixation (3.96 versus 1.59 μmol N m^{-3} day^{-1}), which indicates the increase of physical input of new N was the dominant factor for the PP increase in this layer. The inhibited N_2 fixation and therefore the enhanced N recycling efficiency were a less important mechanism.

The semilabile DON consumption in 50 – 70 m increased by 0.46 μmol N m^{-3} day^{-1} (**Figure 20**), similar to the 0 – 20 m layer. But it was less important in 50 – 70 m compared to the increase of physical new N input. In addition, the biological production of semilabile DON actually increased by 0.68, which is even 50% more than the increase of semilabile DON consumption. That is, the increased N supply from semilabile DON in this layer completely derived from recycled N from biological production.

Thus for 50 – 70 m, the physical input of new N was the major factor controlling the interannual PP increase, and the inhibited N_2 fixation and the enhanced N recycling efficiency are contributing mechanisms, although there were less important. But the N supply from semilabile DON did not contribute to the PP increase.

4. Discussion

4.1. Model explanations for the interannual primary production increase

As more observatories have been set up in the open ocean and more time-series of observations are routinely available, it becomes possible to study the long-term change of oceanic ecosystems and their relationships to the global system. This study is an example to link time-series data to modeling. The data assimilation procedure objectively optimized some of the important, but not determinable, model parameters and successfully minimized the differences between time-series of observations and modeled results at Station ALOHA. The optimized model results showed good fit to most of the 17 types of observations. The model had difficulties fitting several observation types, but this shortcoming could provide useful information to evaluate the limitations of the model structure. Nevertheless, the robustness of the model was supported by the simultaneous good fits to most of the major biogeochemical observations over ~ 150 profiles in 15 years.

As I discussed above, my computation of depth-integrated annual average PP at Station ALOHA from the HOT dataset indicated a difference between the beginning and end of the record of 8.4% (2005 minus 1991). The Model-I linear regression identified the uptrend as significant and predicted the increase rate as ~ 1% per year, i.e., totally ~

14% over this period 1991 – 2005. These growth rates are lower than those reported by Corno et al. (2007) (CN07 hereafter), which estimated ~ 50% increase of PP over 1989 – 2004 from the same HOT dataset. This study linearly interpolated PP over depth to compute depth-integrated PP for each observed profile. The interpolation method used by CN07 appears to be similar to this study because in the first 4 periods defined by CN07 (covering 1989 – 1998) the computed average depth-integrated PP rates are close in both studies (**Table 2**). However, CN07 obtained a much higher depth-integrated PP rate in the period of 1999 – 2004 than this study. The depth-integrated PP of this study was estimated from the original PP data (light 12 hrs plus dark 12 hours) downloaded from HOT data site (http://hahana.soest.hawaii.edu/hot/hot-dogs/index.html). In 1989 – 1990, the measured PP rates were much lower than other years, and the low rates never occurred again in the following period. This is the reason my study did not include 1989 and 1990. Thus without including 1989 and 1990, the total increase of PP in 1991 – 2004 could be ~ 27% from CN07. Furthermore, if the CN07's estimate for 1999 – 2004 was replaced by the estimate of this study, the total increase of PP in 1991 - 2004 could be ~ 15%.

The model after optimization was able to simulate the decadal increase of PP. The results showed that the increase was mainly realized by normal phytoplankton and unicellular diazotrophs. The simulated PP by *Trichodesmium* spp. did not show significant increase over years. The increase in PP occurred mainly in the surface 80 m and showed different features in the surface 50 m and 50 – 80 m. The increase in the surface 50 m was vertically uniform and represented the pattern in the mixed layer. The 50 – 80 m layer, especially 50 – 70 m, had the highest rate of increase. I found the PP

increase in the surface 20 m was controlled by two mechanisms: First, the N_2 fixation was inhibited by the increased physical supply of DIN from depth, and this transformation between the two forms of new N resulted in an enhanced N recycling efficiency. Second, the production and consumption of semilabile DON was not coupled on the time scale of this study, and net semilabile DON consumption acts as another new N source. In the 50 – 70 m layer, the increased physical supply of DIN was the dominant the controlling mechanism, and the reduction of N_2 fixation and enhanced N recycling efficiency were less important. The net N consumption from semilabile DON did not contribute to the PP increase in 50 – 70 m, because the increase of its consumption rate was lower than the increase of its production rate from biological species, which was a recycled N form.

4.2. Increase physical DIN input

The increased physical input of the DIN was one of the main reasons for the interannual increase of PP in the both layers. Physical transport in the model was driven by three mechanisms. First, vertical **advection** was controlled by vertical velocity. Second, vertical **diffusion** was controlled by vertical eddy diffusivity, which was set to a constant value of $1.1\times10^{-4}m^2s^{-1}$ at the mixed layer depth (MLD) and exponentially decreased below MLD. This diffusivity did not apply in the mixed layer. Instead, the third mechanism of physical transport was the completely **mixing** of all the materials inside the mixed layer at every time step (see details in Chapter 2). Thus both the vertical diffusion and mixing of DIN depend on the MLD. In 0 – 20 m, most of the physical input of DIN was realized by mixing, whose annual average increased by 1.09 μmol N m^{-3} day^{-1} from 1991 to 2005. The physical input of DIN by advection and diffusion was

several orders less than that by mixing in this layer. Thus the deepening mixed layer was expected to control the vertical DIN input to the surface 20 m. In 50 – 70 m, the increase of physical input of DIN was realized by all the three mechanisms: 0.36 increase by advection, 2.53 increase by diffusion and 1.07 increase by mixing. The DIN gradient increased by 7.8 μmol N m^{-3} m^{-1} at 70 m depth and by 0.7μmol N m^{-3} m^{-1} at 50 m depth, while the vertical diffusivity did not change significantly at these two depths. Thus the increase of diffusion was generated by the increased DIN gradient, which could resulted from deepened MLD. Thus most of the increase, i.e., diffusion plus mixing, depended on the MLD. The MLD used to force the model was derived from the linear interpolation of the monthly HOT observations, using the criteria of a temperature change of 0.5ºC from the sea surface (see details in Chapter 2). The annual average of the MLD had showed a trend of deepening from 1991 to 2002, and shoaled from 2003 to 2005 (**Figure 21b**). Although the winter MLD in 2003 – 2005 was shallower than the previous period, the summer MLD in these years were deeper (**Figure 21a**). Thus the actual effect in 2003 – 2005 would not be as significant as shown by the annual average MLD.

This result was consistent with one of the conclusions by Corno et al. (2007). They indicated that one of the reasons for the increased PP was the decreased physical stability and increased MLD at Station ALOHA, related to El Niño/ Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO) variations. Note that some of previous works indicated N_2 fixation at Station ALOHA was developing because of the intensified stability (e.g. Karl et al. 1997). But those were the conclusions on time scale of a season to several years. But on the decadal time scale, the physical stability was decreasing at Station ALOHA.

4.3. N_2 fixation and the related Low N Recycling Efficiency

The modeled results showed that the new N sources from N_2 fixation were less important than physical DIN input in terms of their enhancement for PP, because the new N from the major N_2 fixer in the model, *Trichodesmium* spp., had a lower recycling efficiency than N from other phytoplankton. It has been recognized that *Trichodesmium* spp. lack major grazers because some of them are able to produce a toxin (Hawser et al. 1992) and majority of their cells aggregate as colonies (Karl et al. 2002). Although a specialized group of harpacticoid copepods appear adapted to the toxin and are capable of directly grazing *Trichodesmium* spp. (O'Neil & Roman 1994), their quantitative role in the consumption of *Trichodesmium* spp. is unknown. The extracellular release of dissolved organic and inorganic N (Glibert & Bronk 1994, Mulholland et al. 2004) is probably the major pathways for *Trichodesmium* spp. to provide their newly fixed N to the ecosystem (Capone et al. 1997, Mulholland 2007). Viral lysis may be another pathway for *Trichodesmium* spp. to provide the fixed N (Hewson et al. 2004). The extracellular release of dissolved N by *Trichodesmium* spp. is represented by the model as 30% of newly fixed N per day. The viral lysis is not represented by the model explicitly but instead is implicitly included in the passive release of labile DOM by *Trichodesmium* spp., which is totally 2% of their biomass per day and is comparable with the observed 1.65± 0.99% by Hewson et al. (2004). The major reason for the low N recycling efficiency through *Trichodesmium* spp. in the model was that they were only grazed by metazoan zooplankton (e.g., copepods), consistent with observations, while other phytoplankton were grazed by protozoan zooplankton. Thus the conclusion that the N recycling efficiency through *Trichodesmium* spp. is lower depends on validity of two

assumptions: (1) metazoa have lower recycling efficiency than protozoa; and (2) more normal phytoplankton are grazed by protozoa than by metazoa. Regarding the first assumption, metazoa tends to produce larger fecal pellets than protozoa and thus generate more export. Although the model did not have different sizes of detritus, lower detritus production rate were assigned to protozoa than to metazoa in order to represent a smaller size of detritus that stayed in the surface ocean longer and was more likely to be recycled before sinking out of the model domain. Regarding the second assumption, small-size picophytoplankton, like *Prochlorococcus* and *Synechococcus*, dominate the phytoplankton assemblage at Station ALOHA (Campbell 1994). They are too small to be consumed directly by metazoa (Reckermann & Veldhuis 1997). All these points support the idea that the pathway through *Trichodesmium* spp. has lower N recycling efficiency than other phytoplankton.

Another N_2-fixing group in the model, unicellular diazotrophs, had even higher N recycling efficiency than normal phytoplankton. This did not impact my conclusion of low N recycling efficiency for the overall N_2-fixing group in the model, as the major N_2 fixation in the model was realized by *Trichodesmium* spp. However, it was reported that the N_2 fixation rate by unicellular diazotrophs could equal to *Trichodesmium* spp. (Montoya et al. 2004). In the model, the normal phytoplankton and unicellular diazotrophs were treated equally in terms of predation by protozoa, and the "extra" extracellular release of labile DON and ammonium helped the unicellular diazotrophs to have higher recycling efficiency. But in reality the size of unicellular diazotrophs is normally larger than the majority of normal phytoplankton, *Prochlorococcus* (Zehr et al. 2001). The predation rates on these two groups can be different, which, however, is still

unclear with current limited knowledge on unicellular diazotrophs. If N_2 fixed by unicellular diazotrophs and *Trichodesmium* spp. was about equal and also the N recycling efficiencies through normal phytoplankton and unicellular diazotrophs were close, my conclusion about low N recycling efficiency for N_2 fixation would be still valid.

5. Conclusion

This study successfully applied data assimilation technology to a marine ecosystem model and optimized the model to fit the 15-year major biogeochemical observations at Station ALOHA. Data assimilation objectively tested the robustness of the model. The model simulated an interannual increase of primary production at Station ALOHA. The results indicated that the increased mixed layer depth and therefore the increased physical input of DIN was one of the major reasons for the increase of primary production. The increased physical DIN input inhibited N_2 fixation. The model suggested that diazotrophs had lower nitrogen recycling efficiencies and the replacement of new nitrogen from N_2 fixation by physical input could enhance primary production even the total amount of new N did not change. The semilabile DON acts as another new form of N on the time scale of the study period in surface 20 m, which was another important mechanism supporting the interannual increase of PP.

Although significant changes have been observed and also successfully simulated by the model, the study period of 15 years is still too short to show clear trends. The seasonal and interannual changes showed the possibility for these trends to reverse. Longer study is needed to confirm whether these changes are definitive.

References

Campbell L, H. A. Nolla, and D. Vaulot (1994) The importance of Prochlorococcus to community structure in the central North pacific Ocean. Limnol. Oceanogr. 39:954-961

Capone DG, Ferrier MD, Carpenter EJ (1994) Amino-Acid Cycling in Colonies of the Planktonic Marine Cyanobacterium Trichodesmium-Thiebautii. Appl. Envir. Microbiol. 60:3989-3995

Capone DG, Zehr JP, Paerl HW, Bergman B, Carpenter EJ (1997) Trichodesmium, a globally significant cyanobacterium. Science 276:1221-1229

Carlson CA (2002) Production and removal processes. In: Hansell DA, Carlson CA (eds) Biogeochemistry of marine dissolved organic matter. Elsevier Science, USA, New York, p 91-151

Casciotti KL, Trull TW, Glover DM, Davies D (2008) Constraints on nitrogen cycling at the subtropical North Pacific Station ALOHA from isotopic measurements of nitrate and particulate nitrogen. Deep-Sea Res. II 55:1661-1672

Church MJ, Ducklow HW (2002) Multiyear increases in dissolved organic matter inventories at Station ALOHA in the North Pacific Subtropical Gyre. Limnol Oceanogr 47:1-10

Church MJ, Ducklow HW, Letelier RM, Karl DM (2006) Temporal and vertical dynamics in picoplankton photoheterotrophic production in the subtropical North Pacific Ocean. Aquat. Microb. Ecol. 45:41-53

Church MJ, Mahaffey C, Letelier RM, Lukas R, Zehr JP, Karl DM (2009) Nitrogen fixation in the North Pacific Subtropical Gyre: The role of mesoscale physical

forcing. Global Biogeo. Cycles in press

Corno G, Karl DM, Church MJ, Letelier RM, Lukas R, Bidigare RR, Abbott MR (2007) Impact of climate forcing on ecosystem processes in the North Pacific Subtropical Gyre. Journal of Geophysical Research-Oceans 112:Artn C04021, Doi 10.1029/2006jc003730

Doney SC, Ducklow HW (2006) A decade of synthesis and modeling in the US Joint Global Ocean Flux Study. Deep-Sea Res. II 53:451-458

Ducklow HW, Carlson CA (1992) Oceanic bacterial productivity. Adv. Microb. Ecol. 12:113-181

Dugdale RC, Goering JJ (1967) Uptake of new and regenerated forms of nitrogen in primary production. Limnol. Oceanogr. 12:196-206

Friedrichs MAM, Dusenberry JA, Anderson LA, Armstrong RA, Chai F, Christian JR, Doney SC, Dunne J, Fujii M, Hood R, McGillicuddy DJ, Moore JK, Schartau M, Spitz YH, Wiggert JD (2007) Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. J. Geophys. Res. 112:C08001, Doi 10.1029/2006jc003852

Glibert PM, Bronk DA (1994) Release of dissolved organic nitrogen by marine diazotrophic cyanobacteria, Trichodesmium spp. Appl. Envir. Microbiol. 60:3996-4000

Hawser SP, Oneil JM, Roman MR, Codd GA (1992) Toxicity of Blooms of the Cyanobacterium-Trichodesmium to Zooplankton. J. Appl. Phycol. 4:79-86

Hedges JI (1992) Global biogeochemical cycles: progress and problems. Mar. Chem. 39:67-93

Hewson I, Govil SR, Capone DG, Carpenter EJ, Fuhrman JA (2004) Evidence of *Trichodesmium* viral lysis and potential significance for biogeochemical cycling in the oligotrophic ocean. Aquat. Microb. Ecol. 36:1-8

Karl D, Michaels A, Bergman B, Capone D, Carpenter E, Letelier R, Lipschitz F, Paerl H, Sigman D, Stal L (2002) Dinitrogen fixation in the world's oceans. Biogeochemistry 57/58:47-98

Karl DM, Bjoerkman KM, Dore JE, Fujieki L, Hebel DV, Houlihan T, Letelier RM, Tupas LM (2001a) Ecological nitrogen-to-phosphorus stoichiometry at station ALOHA. Deep-Sea Res. II 48:1529-1566

Karl DM, Dore JE, Lukas R, Michaels AF, Bates NR, Knap A (2001b) Building the long-term picture: the U.S. JGOFS time-series programs. Oceanography 14:6-17

Karl DM, Letelier R, Tupas L, Dore J, Christian J, Hebel D (1997) The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. Nature 388:533-538

Karl DM, Letelier RM (2008) Nitrogen fixation-enhanced carbon sequestration in low nitrate, low chlorophyll seascapes. Marine Ecology-Progress Series 364:257-268

Kirchman DL (2000) Uptake and regeneration of inorganic nutrients by marine heterotrophic bacteria. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, Inc., New York, p 261-288

Lawson LM, Spitz YH, Hofmann EE, Long RB (1995) A data assimilation technique applied to a predator-prey model. Bulletin of Mathematical Biology 57:593-617

Montoya JP, Houll CM, Zehr JP, Hansen A, Villareal TA, Capone DG (2004) High rates of N_2 fixation by unicellular diazotrophs in the oligotrophic Pacific Ocean. Nature

430:1027-1032

Mulholland MR (2007) The fate of nitrogen fixed by diazotrophs in the ocean. Biogeosciences 4:37-51

Mulholland MR, Bronk DA, Capone DG (2004) Dinitrogen fixation and release of ammonium and dissolved organic nitrogen by Trichodesmium IMS101. Aquat. Microb. Ecol. 37:85-94

Mulholland MR, Lomas MW (2008) Nitrogen Uptake and Assimilation. In: Capone DG, Bronk DA, Mulholland MR, Carpenter EJ (eds) Nitrogen in the Marine Environment (2nd Edition). Academic Press, Burlington, MA, USA

O'Neil JM, Metzler PM, Glibert PM (1996) Ingestion of $^{15}N_2$-labelled Trichodesmium spp and ammonium regeneration by the harpacticoid copepod Macrosetella gracilis. Mar. Biol. 125:89-96

O'Neil JM, Roman MR (1994) Ingestion of the cyanobacterium Trichodesmium spp. by pelagic harpacticoid copepods Macrosetella, Miracia and Oculosetella. Hydrobiologia 292-293:235-240

Reckermann M, Veldhuis MJW (1997) Trophic interactions between picophytoplankton and micro- and nanozooplankton in the western Arabian Sea during the NE monsoon 1993. Aquat. Microb. Ecol. 12:263-273

Williams PJl (2000) Heterotrophic bacteria and the dynamics of dissolved organic material. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, New York, p 153- 200

Zehr JP, Waterbury JB, Turner PJ, Montoya JP, Omoregie E, Steward GF, Hansen A, Karl DM (2001) Unicellular cyanobacteria fix N_2 in the subtropical North Pacific

Ocean. Nature 412:635-638

Table 1. Default model parameter values, and optimized model parameter values after data assimilation with uncertainty percentage of optimized parameters (if constrained) and gradients of parameters with respect to the cost function after data assimilation. See text for the definition of the gradient. Notes column indicates if the parameter was optimized and well constrained (CS), optimized but not constrained (NC) or not optimized (empty) by data assimilation. Please see Chapter 2 for parameter definitions and units.

Parameter Symbol	Default Value	Optimized Value	Uncertainty	Gradient	Notes
A_E	4000	4000		0.36	
μ_{PHY}	3.5	1.59	11%	0.43	CS
α_{PHY}^{CHL}	2.0	1.70	28%	0.14	CS
β_{PHY}	0.0005	0.0005		2.0	
$v_{REF,PHY}^{N}$	0.3	0.25		-1.5	NS
k_{PHY}^{NH4}	0.01	0.01		-0.078	
k_{PHY}^{NO3}	0.1	0.1		0.87	
$v_{REF,PHY}^{P}$	0.02	0.013		0.72	NS
k_{PHY}^{PO4}	0.0048	0.0048		-0.13	
ζ^{NO3}	2.0	2.0		0.78	
θ	2.0	2.20	20%	0.037	CS
$ex_{PSV,PHY}$	0.05	0.05		2.9	
$ex_{CHO,PHY}$	0.05	0.05		2.3	
pom_{PHY}	0.05	0.0354	34%	0.21	CS
μ_{TR}	2.0	2.60	16%	-0.13	CS
α_{TR}^{CHL}	0.25	0.20		1.5	NS
β_{TR}	0.002	0.002		0.64	
$v_{REF,TR}^{N}$	0.4	0.25		0.45	NS
k_{TR}^{NH4}	0.02	0.02		-0.19	
k_{TR}^{NO3}	0.2	0.2		0.15	

$v_{REF,TR}^{P}$	0.02	0.020		0.64	
k_{TR}^{PO4}	0.047	0.047		-0.59	
$pick_{TR}^{PO4}$	1.0	1.0		-0.33	
ζ^{NFIX}	3.0	3.0		-0.077	
$ex_{PSV,TR}$	0.02	0.02		0.059	
$ex_{NFIX,TR}$	0.30	0.30		0.15	
$ex_{CHO,TR}$	0.05	0.05		0.27	
pom_{TR}	0.1	0.080		0.25	NS
μ_{UN}	0.5	0.5		-2.1	
α_{UN}^{CHL}	1.0	1.0		-0.89	
k_{DOM}	0.5	0.5		0.63	
$b_{LABI,N}$	15	10		0	NS
$b_{LABI,P}$	15	9.95	3.7%	1.4	CS
μ_{BA}	2.0	2.0		-0.54	
b_{RESP}	0.28	0.28		0.35	
$ex_{ADJ,BA}$	2.0	2.0		0.00	
$remi_{BA}$	3.0	6.0		0.37	NS
$ex_{REFR,BA}$	0.01	0.01		0.57	
$f_{SLCT,BA}$	0.2	0.2		-0.32	
$resp_{BA}^{B}$	0.01	0.01		0.46	
$r_{\min,BA}^{A}$	0.35	0.35		-0.46	
$r_{\max,BA}^{A}$	0.80	0.621	5.9%	0.34	CS
$mort_{BA}$	0.02	0.049		3.1	NS
μ_{PRT}	4.0	4.0		-6.6	
g_{PHY}	0.67	1.23	15%	0.53	CS
g_{BA}	2.0	1.93	21%	-0.19	CS
ex_{PRT}	0.15	0.15		3.0	
$f_{ex,PRT}$	0.75	0.80		1.2	NS
$resp_{PRT}^{B}$	0.01	0.01		-1.7	
$resp_{PRT}^{A}$	0.67	0.787	2.1%	1.4	CS
$ex_{ADJ,PRT}$	2.0	2.0		0.017	
$remi_{PRT}$	4.7	4.7		-0.66	

pom_{PRT}	0.027	0.027		0.93	
μ_{MZ}	1.0	1.3		-0.32	NS
g_{PRT}	0.65	1.13	32%	-0.18	CS
g_{TR}	0.25	0.25		1.9	
ex_{MZ}	0.20	0.20		-0.26	
$f_{ex,MZ}$	0.75	0.75		-1.5	
$resp_{MZ}^{B}$	0.05	0.030		-0.25	NS
$resp_{MZ}^{A}$	0.25	0.217		-1.9	NS
$ex_{ADJ,MZ}$	2.0	2.0		0.00	
$remi_{MZ}$	4.0	4.0		-0.0078	
pom_{MZ}	0.15	0.15		0.13	
$ex_{REFR,MZ}$	0.002	0.002		-0.021	
$remv_{MZ}$	0.5	0.81		-1.8	NS
$f_{SDOM,HZ}$	0.1	0.10		0.70	
$f_{POM,HZ}$	0.14	0.14		0.68	
$ex_{REFR,SDOM}$	0.0001	0.0001		-0.83	
$q_{N,REFR}^{C}$	0.058	0.058		-2.1	
$q_{P,REFR}^{C}$	0.00076	0.00076		0.026	
$q_{N,POM}^{C}$	0.12	0.12		1.3	
$q_{P,POM}^{C}$	0.0045	0.0045		-0.57	
r_{ntrf}	0.1	0.1		0.60	
prf_{N}	1.1	1.1		0.24	
prf_{P}	4.0	7.4		-0.21	NS
wnsv	2.0	5.02	28%	0.13	CS
diss	0.02	0.011		0.75	NS

Table 2. Computed average depth-integrated primary production and one standard deviation (mmol C m^{-2} d^{-1}) at Station ALOHA in 5 periods from the work of Corno et al. (2007) and this study.

Year	1989 - 1990	1991 - 1994	1995 - 1996	1997 - 1998	1999 - 2004
Corno	33±4	41±1	41±3	42±2	52±2
This study	34.8±6.1	41.6±3.3	44.3±5.0	41.0±3.9	46.9±3.4

Figure Legends

Figure 1. Comparison of each component (data type) to the cost function, J, before and after data assimilation. Y axis is in logarithmic scale. **NO3**: Nitrate; **PO4**: phosphate; **MZc**: metazoan zooplankton carbon biomass; **PHYn**: phytoplankton nitrogen biomass; **CHL:** Chloropyll *a*; **PrPr**: particulate primary production; **BAc**: heterotrophic bacterial carbon biomass; **BPr**: heterotrophic bacterial production; **sDOC, sDON** & **sDOP**: semilabile dissolved organic carbon, nitrogen and phosphorus; **POC, PON** & **POP**: particulate organic carbon, nitrogen and phosphorus; **STc, STn** & **STp**: sinking particle carbon, nitrogen and phosphorus flux collected by sediment traps at 150m.

Figure 2. Comparison of depth-integrated observations and their model equivalents from optimized solutions. The red points and the blue diamonds are observations and their model equivalents, respectively. The red lines and the blue lines are the 3-point averages of the observations and the model equivalents. The nitrate, phosphate and metazoan zooplankton carbon biomass are compared with their values in the mixed layer only. The sinking particulate carbon, nitrogen and phosphorus fluxes are compared at 150 m depth. Note the model equivalents are derived from subsampled results at same depths and times as observations.

Figure 3. Vertical and temporal averages of modeled carbon (C), nitrogen (N) and phosphorus (P) fluxes (arrows) and stocks (boxes) normalized to C, N and P components of primary production (stocks normalized to amount of primary production during 1 day) at Station ALOHA. Red arrows: microbial loop, green arrows: major food web, blue

arrows: import and export including production of refractory DOM (in dashed blue arrows), yellow arrows: detritus production, dashed black arrows. Primary production is marked on top of each phytoplankton box. 1 unit C = 0.26 mmol C m^{-3} (day^{-1}); 1 unit N = 0.038 mmol C m^{-3} (day^{-1}); 1 unit P = 0.0027 mmol C m^{-3} (day^{-1}). See text for acronyms.

Figure 4. The observed (solid red line) and modeled (solid blue line) depth-integrated primary production averaged in each year in 1991 – 2005. The least-square linear regression lines are plotted for the observations (dashed red line) and the model (dashed blue line). The regression equation, correlation R^2 statistic and probability *p*-value are also noted for the observations (red text) and the model (blue text).

Figure 5. Comparison of observed (upper panel) and modeled (lower panel) primary production profiles in 1991 – 2005. The observations are interpolated using triangle-based linear interpolation.

Figure 6. The anomalies of annual average primary production of (non-diazotrophic) normal phytoplankton (green), *Trichodesmium* spp. (pink) and unicellular diazotrophs (red). The anomalies are shown in absolute amount (a), and are also normalized to their means (b).

Figure 7. Linear-regression predicted annual change rate of the modeled primary production of each layer in 1991 – 2005 for (a) normal phytoplankton and (b) unicellular diazotrophs.

Figure 8. Linear-regression predicted annual change rate of the modeled primary production of each month in 1991 – 2005 for (a) normal phytoplankton and (b) unicellular diazotrophs. Please note that the analysis is applied to depth-integrated primary production, thus has different unit than Figure 7.

Figure 9. In 1991 – 2005 at surface 20 m, the average modeled carbon component of primary production (a), carbon/nitrogen ratio of primary production (c), and nitrogen component of primary production (e) for normal phytoplankton. (b), (d) and (f) are same as (a), (c) and (e), but are for unicellular diazotrophs. The linear regression lines, equations, correlation R^2 statistic and probability *p*-value are also included for each panel.

Figure 10. Annual average photosynthetically active radiation (PAR) of the model forcing in 1991 – 2005. The linear regression lines, equations, correlation R^2 statistic and probability *p*-value are also included.

Figure 11. Annual average of nitrate and ammonium in 0 – 20 m in 1991 – 2005.

Figure 12. The net change of annually averaged flux rates (μmol m^{-3} day^{-1}) in the surface 20 m for 1991 and 2005: nitrate and ammonium related fluxes. The positive sign indicates the flux is higher in 2005 than in 1991, and the negative sign indicates it is lower in 2005 than in 1991. See text for acronyms.

Figure 13. Annual average of the physical transport of dissolved organic nitrogen (DIN, nitrate and ammonium) (upper panel) and the rate of N_2 fixation (lower panel) in 0 – 20 m.

Figure 14. Nitrogen recycling efficiency through (a) normal phytoplankton (PHY), (b) unicellular diaztrophs (UN) and (c) *Trichodesmium* spp. (TR). The figure shows the amount of nitrogen that would be exported after 100 units of nitrogen was consumed by different primary producers. The fluxes unrelated to export, e.g., production of labile DOM and ammonium, are not included. The efficiency is estimated from the modeled average status of 0 – 20 m in 1991 – 2005. PRT: protozoan zooplankton, MZ: metazoan zooplankton, DET: detritus, SDOM: semilabile DOM, NO3: nitrate, NH4: ammonium, N2 Fix: N_2 fixation.

Figure 15. Same as Figure 12 but for protozoan and metazoan zooplankton.

Figure 16. Same as Figure 12 but for heterotrophic bacteria and dissolved organic matter. The red dashed circle and number represent total biological production of semilabile DON.

Figure 17. In 0 – 20 m, modeled annual averages of semilabile dissolved organic nitrogen (SDON) production and consumption rate (a), the net flux of SDON, i.e., production minus consumption (c), and the stock of SDON (d). The production rate is the net effect of the biological production and physical exchange (b), and the consumption

rate is the consumption of SDON by heterotrophic bacteria.

Figure 18. In 1991 – 2005 at 50 – 70 m, the average modeled carbon component of primary production (a), carbon/nitrogen ratio of primary production (c), and nitrogen component of primary production (e) for normal phytoplankton. (b), (d) and (f) are same as (a), (c) and (e), but are for unicellular diazotrophs. The linear regression lines, equations, correlation R^2 statistic and probability *p*-value are also included for each panel.

Figure 19. The net change of annually averaged flux rates (μmol m^{-3} day^{-1}) in the 50-70 m for 1991 and 2005: nitrate and ammonium related fluxes. The positive sign indicates the flux is higher in 2005 than in 1991, and the negative sign indicates it is lower in 2005 than in 1991. See text for acronyms.

Figure 20. Same as Figure 19 but for heterotrophic bacteria and dissolved organic matter. The red dashed circle and number represent total biological production of semilabile DON.

Figure 21. The mixed layer depth used to force the model (a) and its annual average (b).

Figure 1

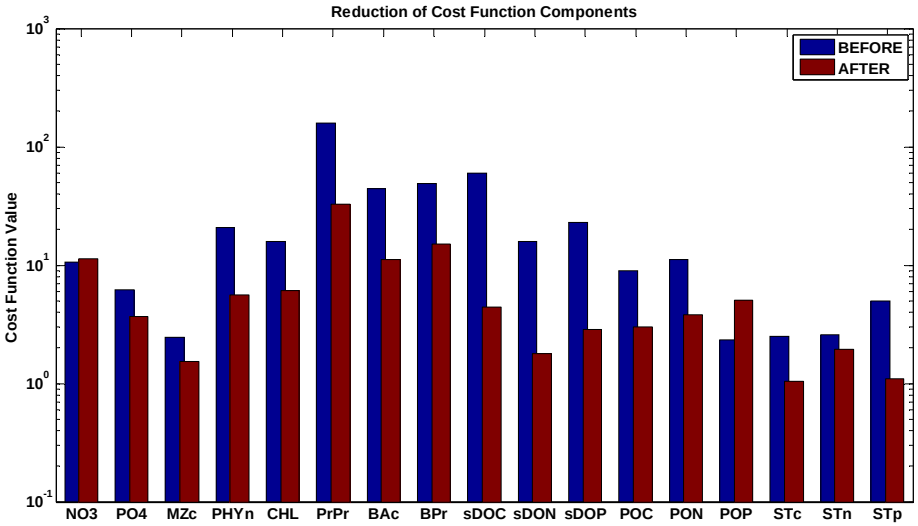
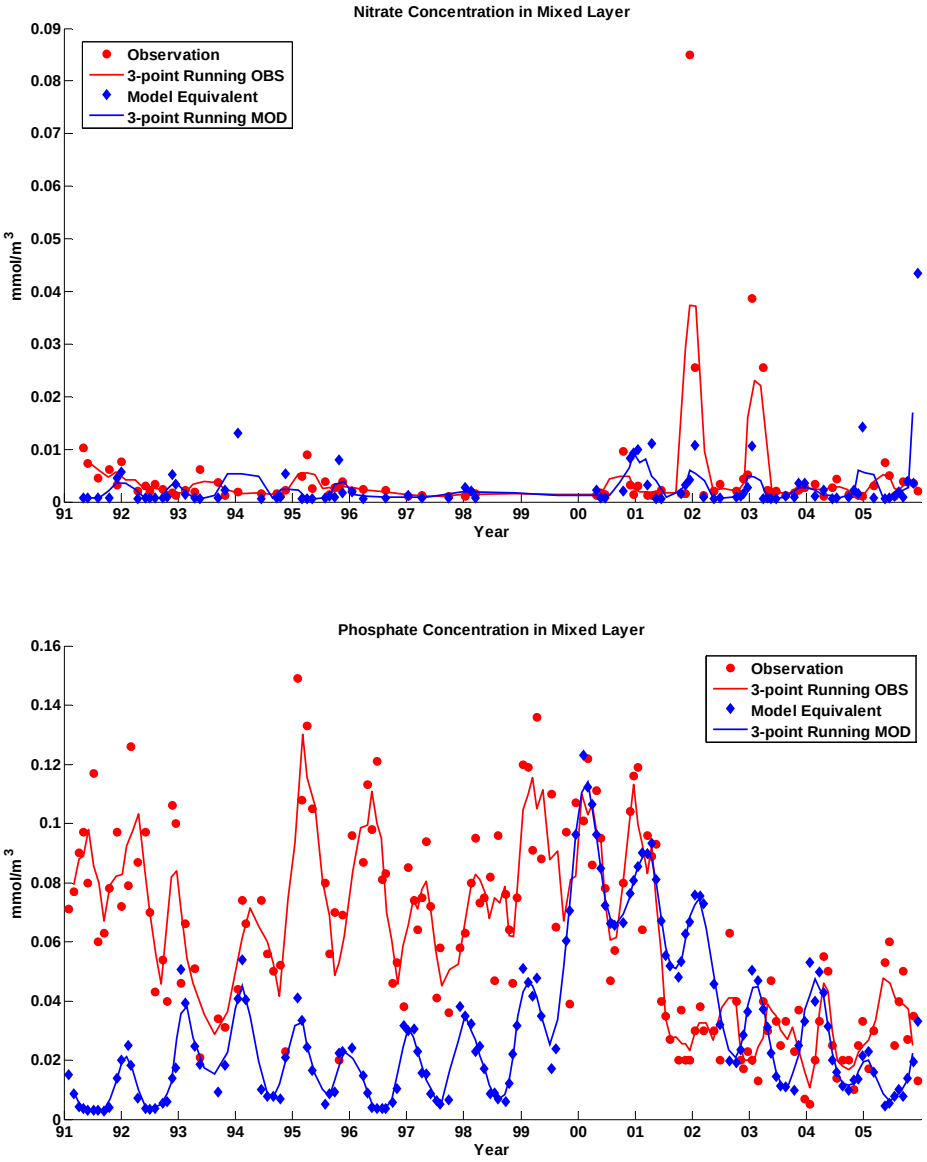

Reduction of Cost Function Components
BEFORE
AFTER
Cost Function Value
10^3
10^2
10^1
10^0
10^-1
NO3
PO4
MZc
PHYn
CHL
PrPr
BAc
BPr
sDOC
sDON
sDOP
POC
PON
POP
STc
STn
STp

Figure 2


Nitrate Concentration in Mixed Layer
Observation
3-point Running OBS
Model Equivalent
3-point Running MOD
mmol/m3
Year
Phosphate Concentration in Mixed Layer
Observation
3-point Running OBS
Model Equivalent
3-point Running MOD
mmol/m3
Year

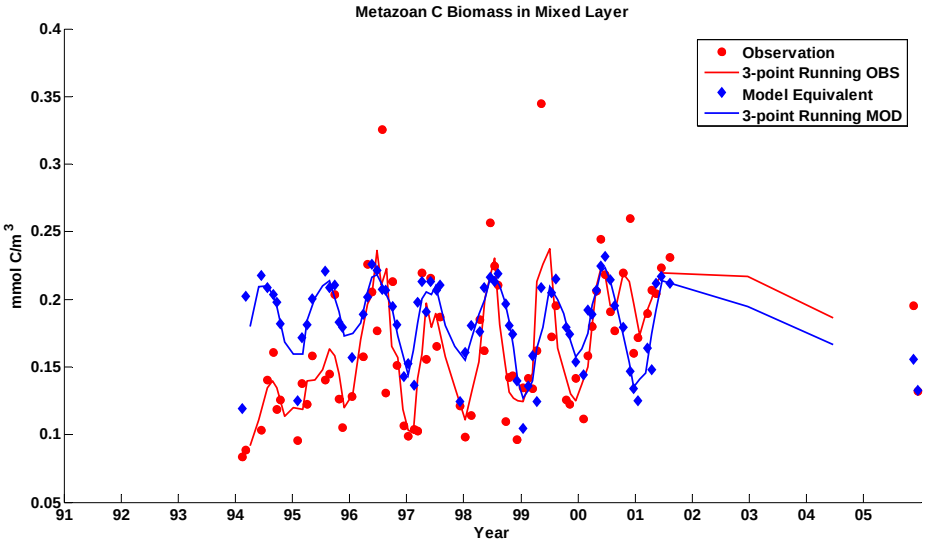
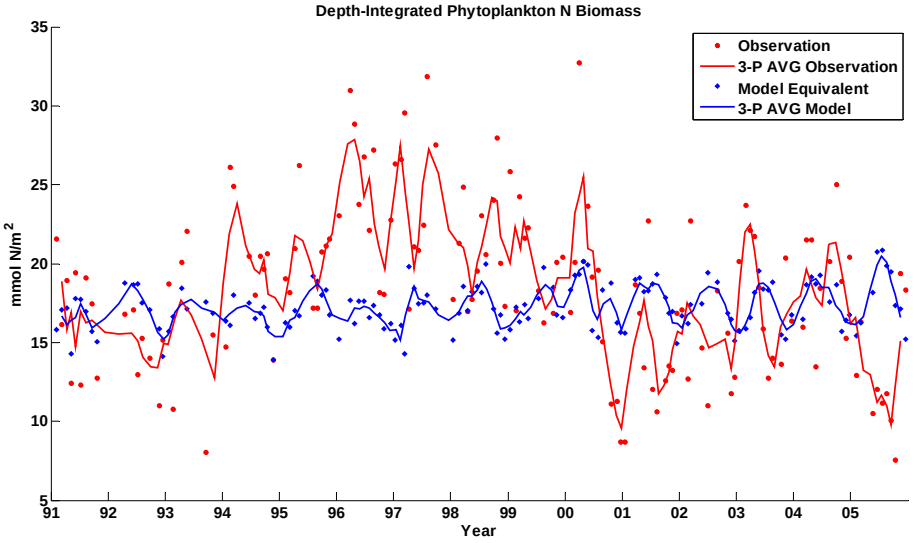


Figure 2 *Continued*



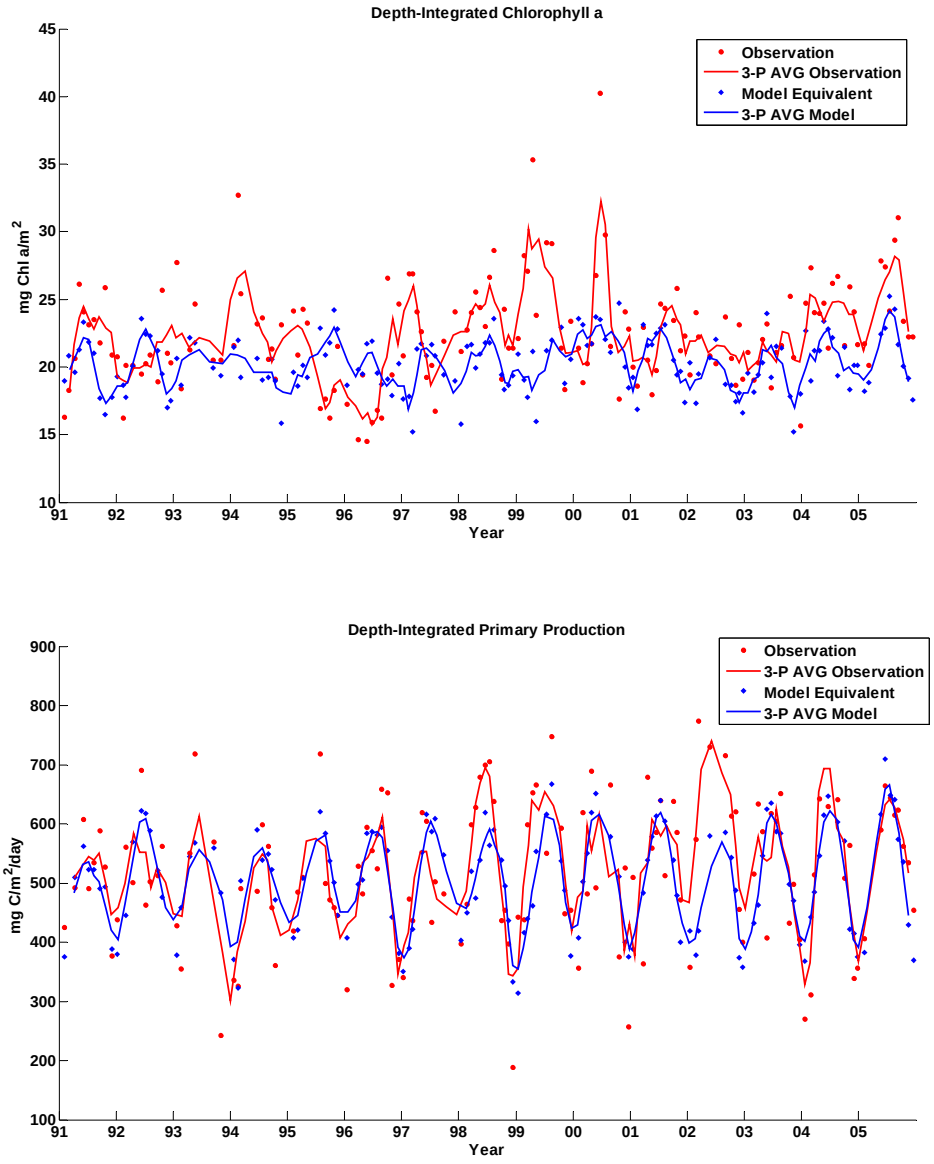
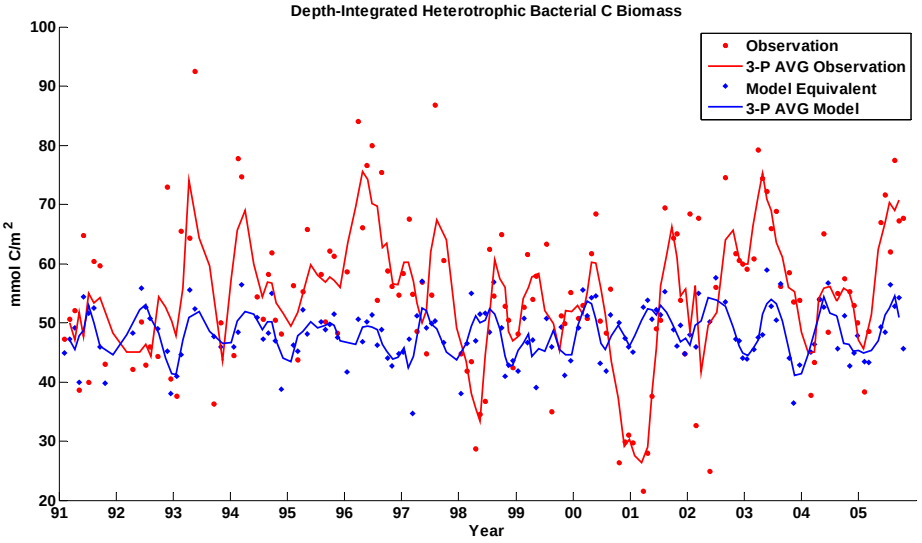
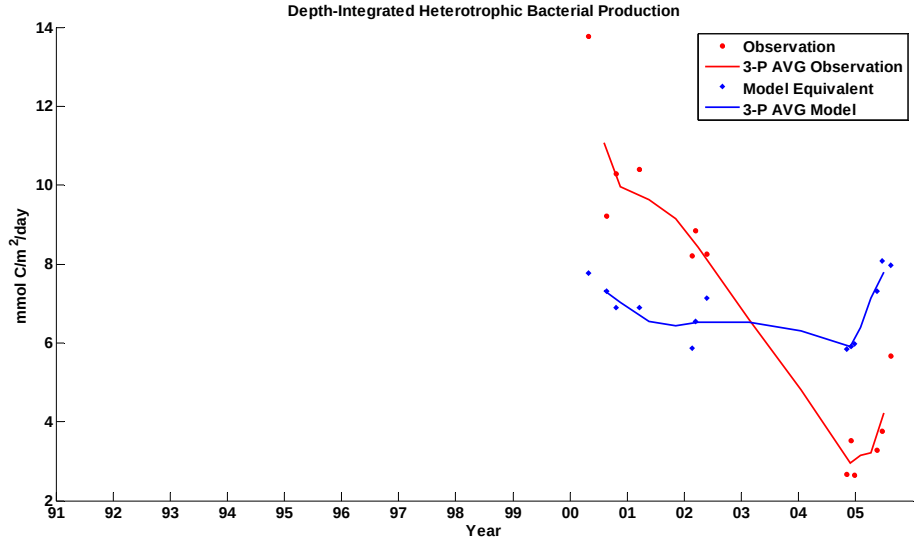


Figure 2 *Continued*


Depth-Integrated Heterotrophic Bacterial C Biomass
Observation
3-P AVG Observation
Model Equivalent
3-P AVG Model
mmol C/m 2
100
90
80
70
60
50
40
30
20
91
92
93
94
95
96
97
98
99
00
01
02
03
04
05
Year


Depth-Integrated Heterotrophic Bacterial Production
Observation
3-P AVG Observation
Model Equivalent
3-P AVG Model
mmol C/m 2/day
14
12
10
8
6
4
2
91
92
93
94
95
96
97
98
99
00
01
02
03
04
05
Year

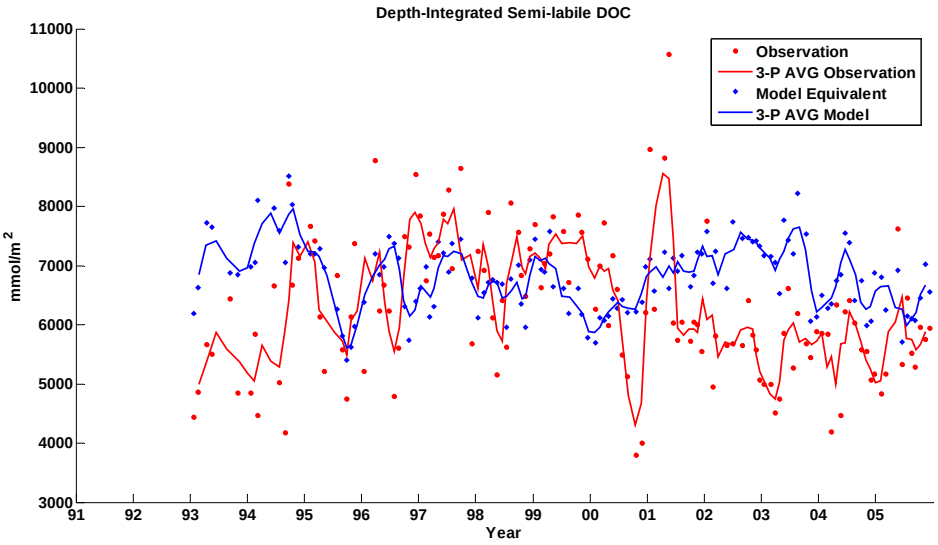
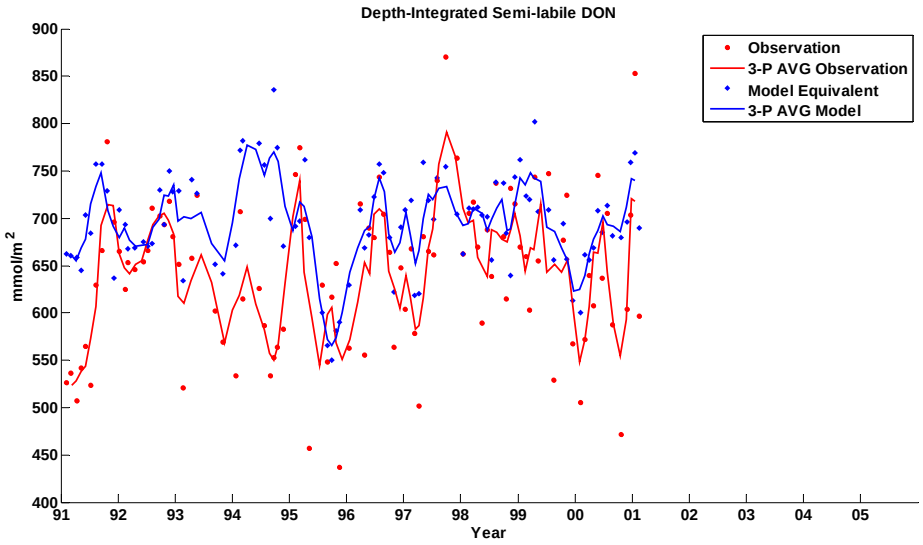


Figure 2 *Continued*



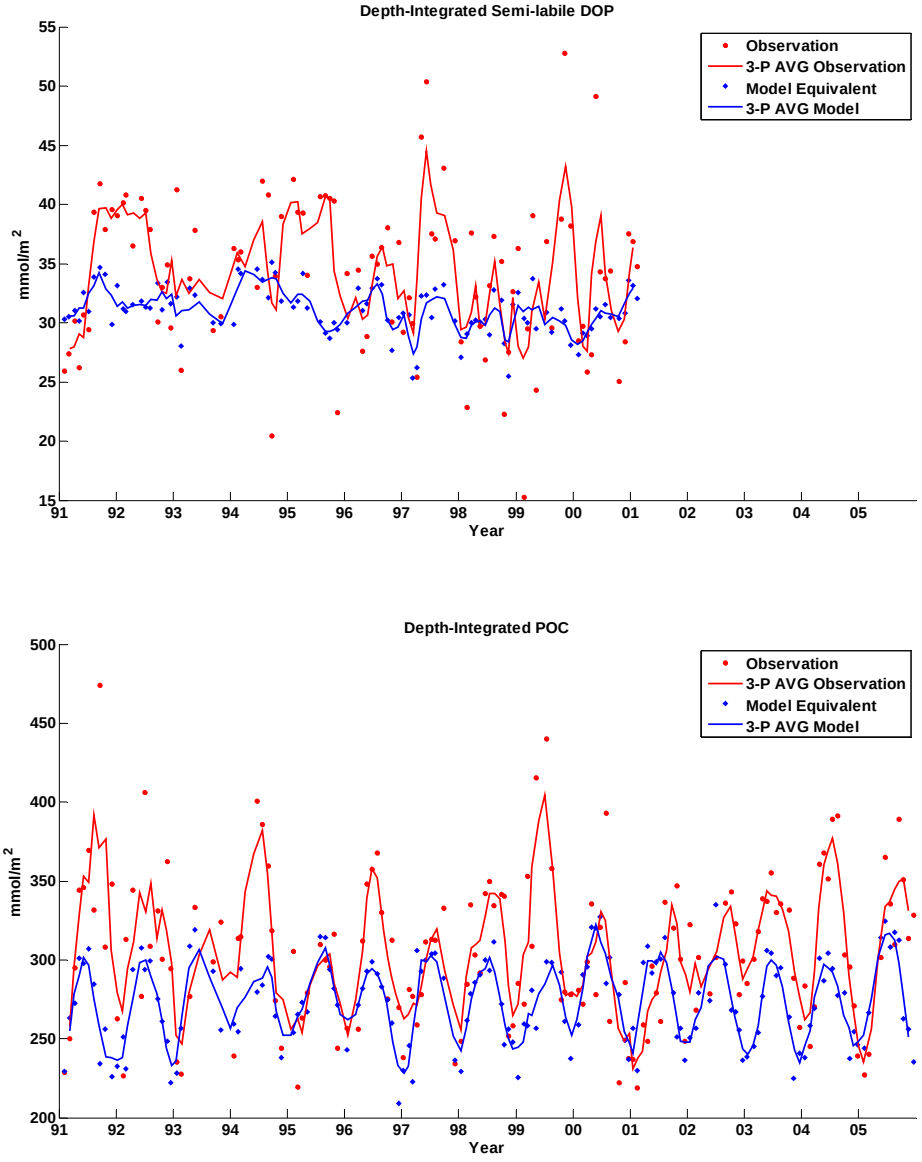
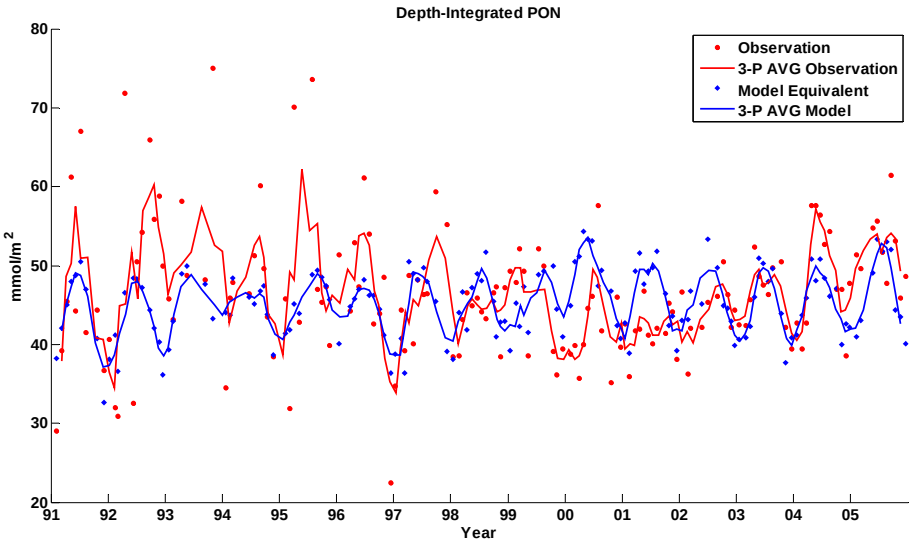
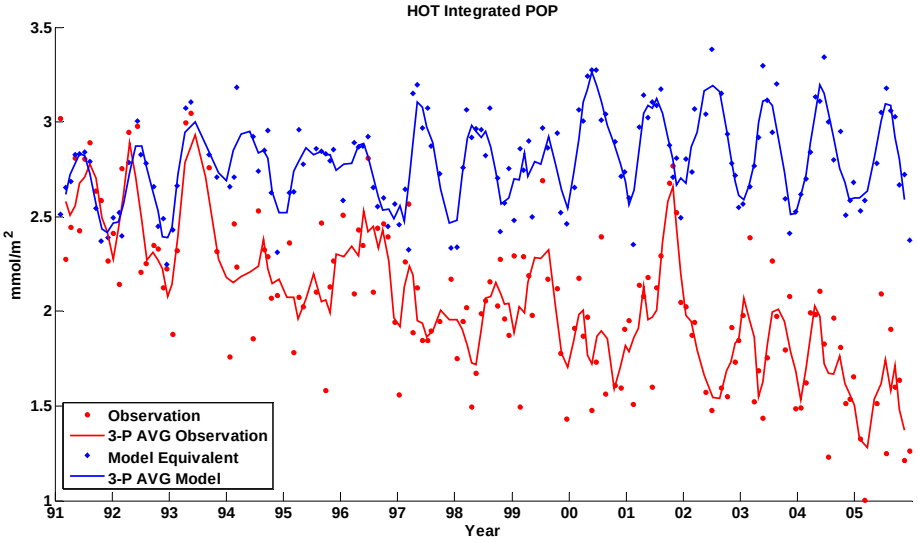


Figure 2 *Continued*


Depth-Integrated PON
Observation
3-P AVG Observation
Model Equivalent
3-P AVG Model
mmol/m 2
Year
80
70
60
50
40
30
20
91
92
93
94
95
96
97
98
99
00
01
02
03
04
05


HOT Integrated POP
Observation
3-P AVG Observation
Model Equivalent
3-P AVG Model
mmol/m 2
Year
3.5
3
2.5
2
1.5
1
91
92
93
94
95
96
97
98
99
00
01
02
03
04
05

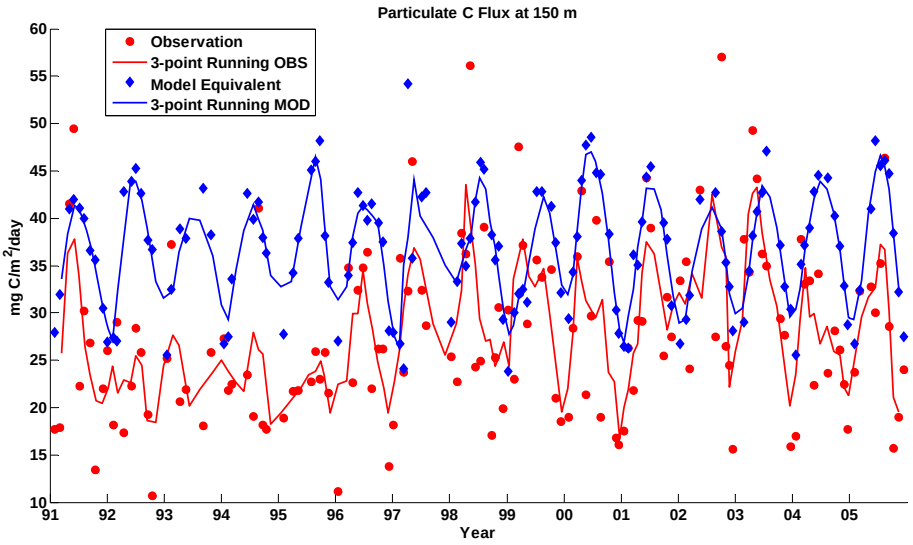
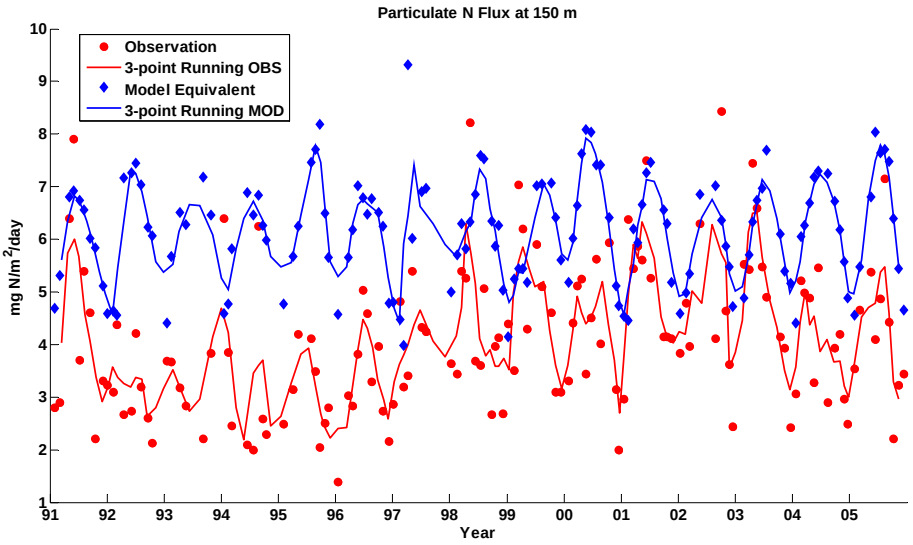


Figure 2 *Continued*



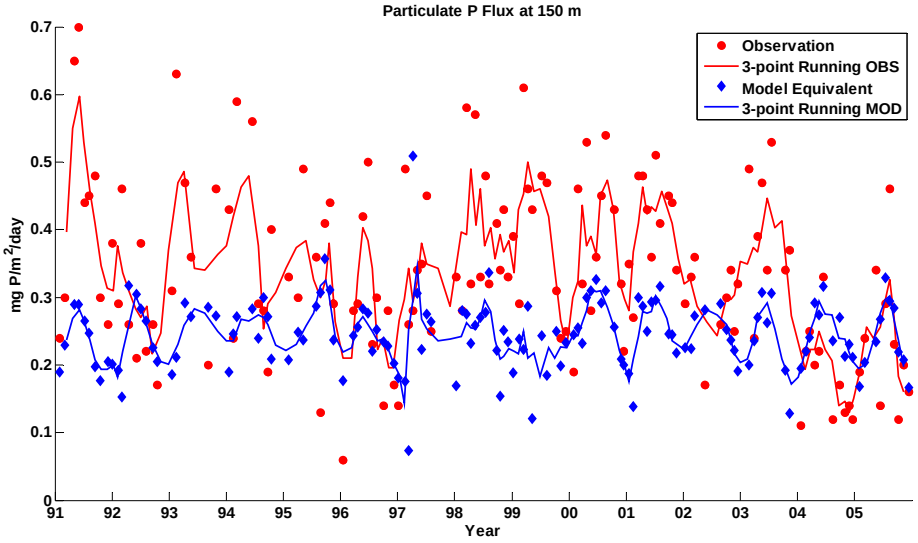
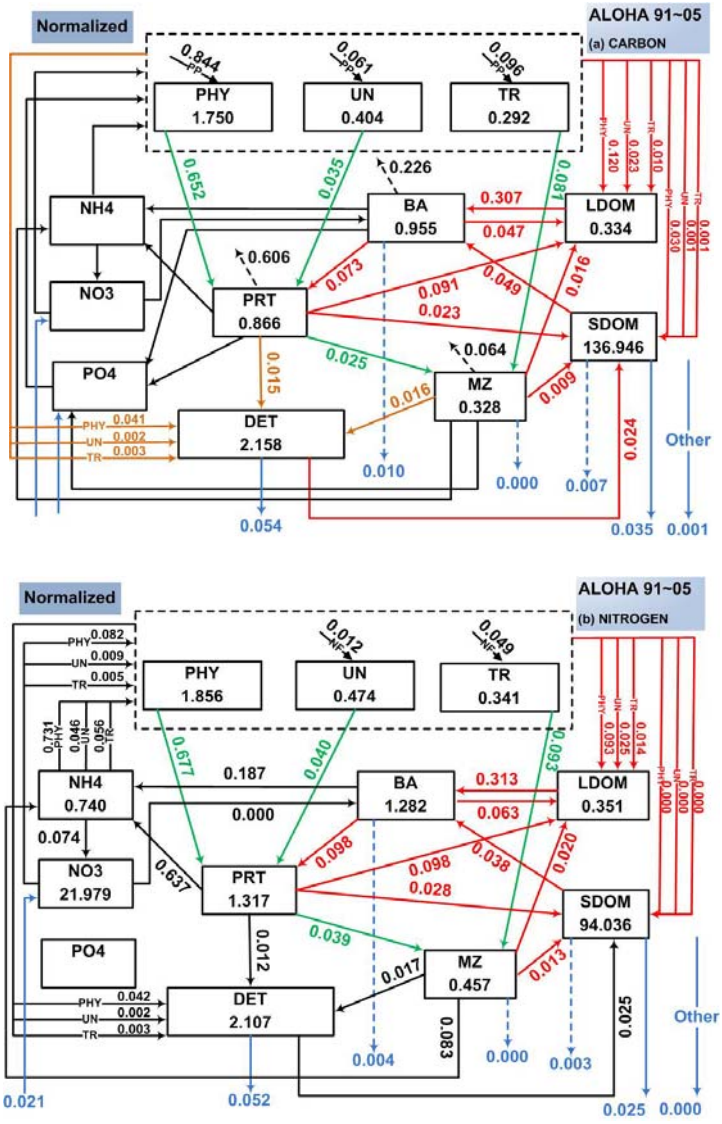

Particulate P Flux at 150 m
Observation
3-point Running OBS
Model Equivalent
3-point Running MOD
mg P/m2/day
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
91
92
93
94
95
96
97
98
99
00
01
02
03
04
05
Year

Figure 3


Normalized
ALOHA 91~05
(a) CARBON
PHY 1.750
UN 0.404
TR 0.292
NH4
NO3
PO4
BA 0.955
LDOM 0.334
PRT 0.866
SDOM 136.946
MZ 0.328
DET 2.158
Other
Normalized
ALOHA 91~05
(b) NITROGEN
PHY 1.856
UN 0.474
TR 0.341
NH4 0.740
NO3 21.979
PO4
BA 1.282
LDOM 0.351
PRT 1.317
SDOM 94.036
MZ 0.457
DET 2.107
Other

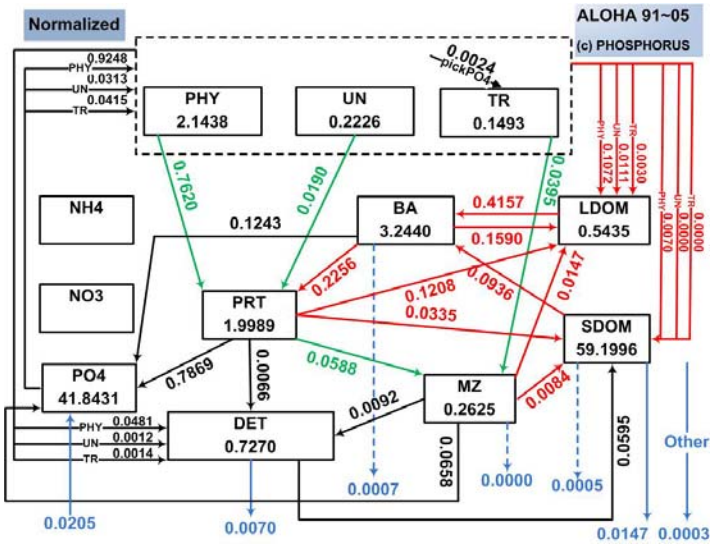

Normalized
ALOHA 91~05
(c) PHOSPHORUS
0.9248
PHY
0.0313
UN
0.0415
TR
PHY
2.1438
UN
0.2226
TR
0.1493
0.0024
pickPO4
0.1072
0.0111
0.0030
0.0070
0.0000
0.0000
NH4
NO3
PO4
41.8431
0.1243
BA
3.2440
LDOM
0.5435
PRT
1.9989
SDOM
59.1996
MZ
0.2625
DET
0.7270
0.7620
0.0190
0.0395
0.4157
0.1590
0.2256
0.1208
0.0335
0.0936
0.0147
0.7869
0.0066
0.0588
0.0092
0.0084
0.0481
0.0012
0.0014
0.0205
0.0070
0.0007
0.0656
0.0000
0.0005
0.0595
0.0147
0.0003
Other

Figure 4

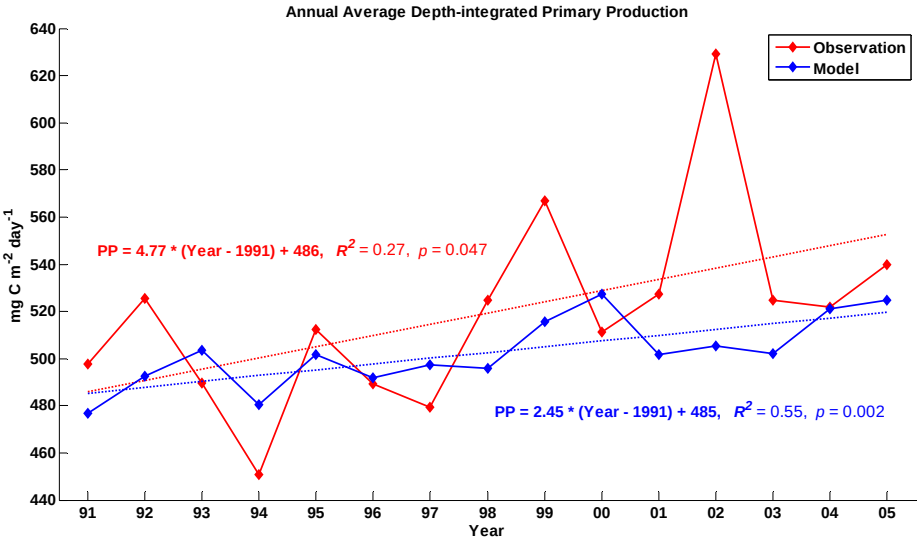

Annual Average Depth-integrated Primary Production
Observation
Model
PP = 4.77 * (Year - 1991) + 486, R2 = 0.27, p = 0.047
PP = 2.45 * (Year - 1991) + 485, R2 = 0.55, p = 0.002
mg C m-2 day-1
640
620
600
580
560
540
520
500
480
460
440
91
92
93
94
95
96
97
98
99
00
01
02
03
04
05
Year

Figure 5

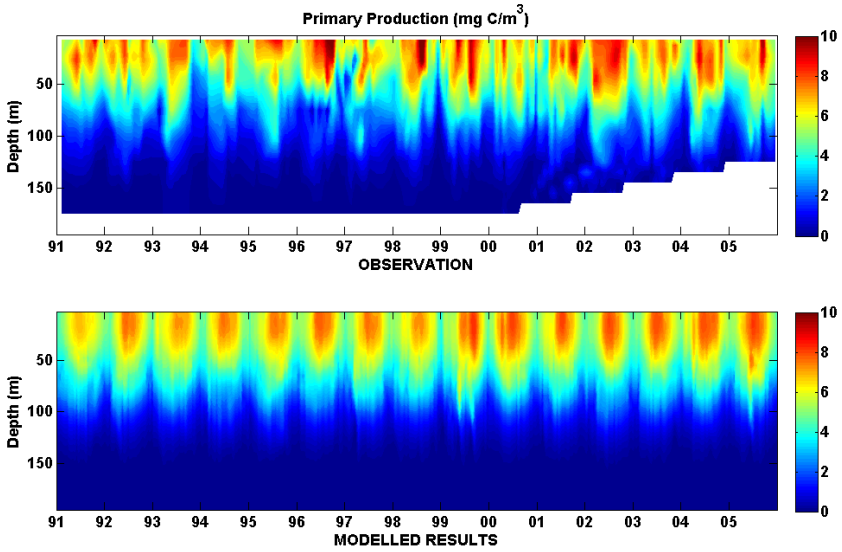

Primary Production (mg C/m3)
Depth (m)
50
100
150
91
92
93
94
95
96
97
98
99
00
01
02
03
04
05
OBSERVATION
10
8
6
4
2
0
Depth (m)
50
100
150
91
92
93
94
95
96
97
98
99
00
01
02
03
04
05
MODELLED RESULTS
10
8
6
4
2
0

Figure 6

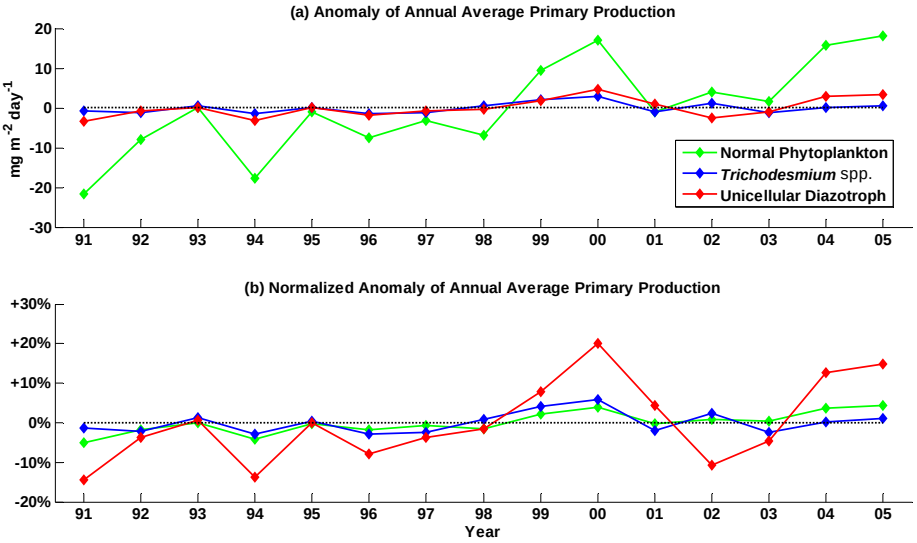

(a) Anomaly of Annual Average Primary Production
mg m-2 day-1
20
10
0
-10
-20
-30
91
92
93
94
95
96
97
98
99
00
01
02
03
04
05
Normal Phytoplankton
Trichodesmium spp.
Unicellular Diazotroph
(b) Normalized Anomaly of Annual Average Primary Production
+30%
+20%
+10%
0%
-10%
-20%
Year

Figure 7

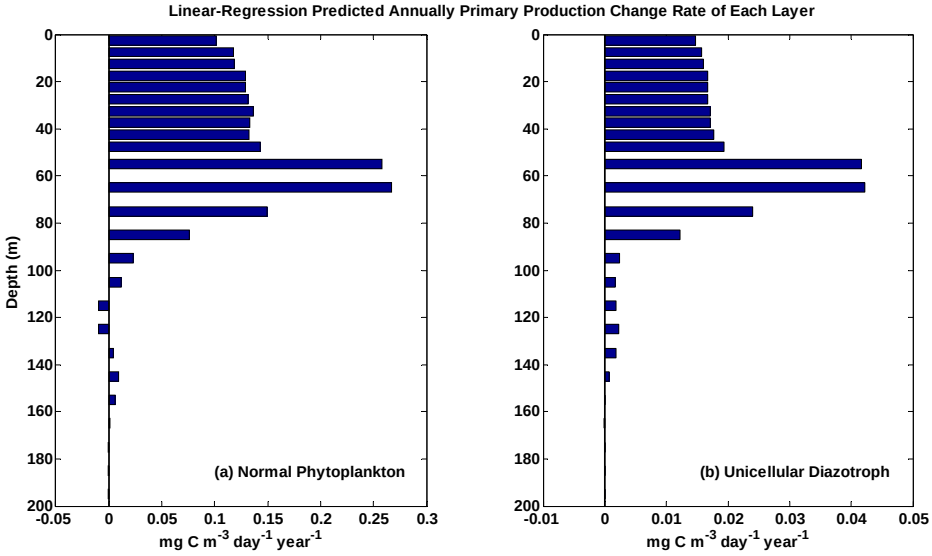

Linear-Regression Predicted Annually Primary Production Change Rate of Each Layer
Depth (m)
0
20
40
60
80
100
120
140
160
180
200
-0.05
0
0.05
0.1
0.15
0.2
0.25
0.3
mg C m-3 day-1 year-1
(a) Normal Phytoplankton
-0.01
0
0.01
0.02
0.03
0.04
0.05
mg C m-3 day-1 year-1
(b) Unicellular Diazotroph

Figure 8

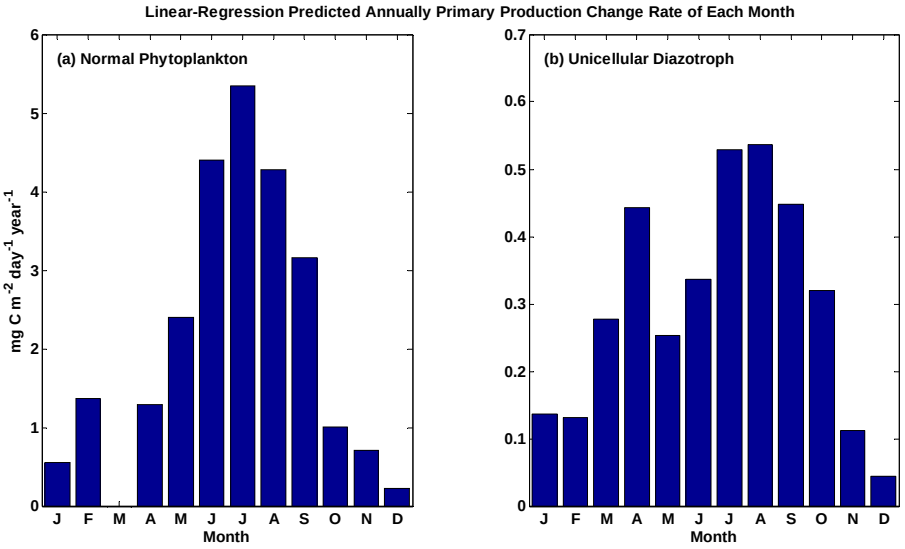

Linear-Regression Predicted Annually Primary Production Change Rate of Each Month
(a) Normal Phytoplankton
(b) Unicellular Diazotroph
mg C m-2 day-1 year-1
Month
J F M A M J J A S O N D
0 1 2 3 4 5 6
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7

Figure 9

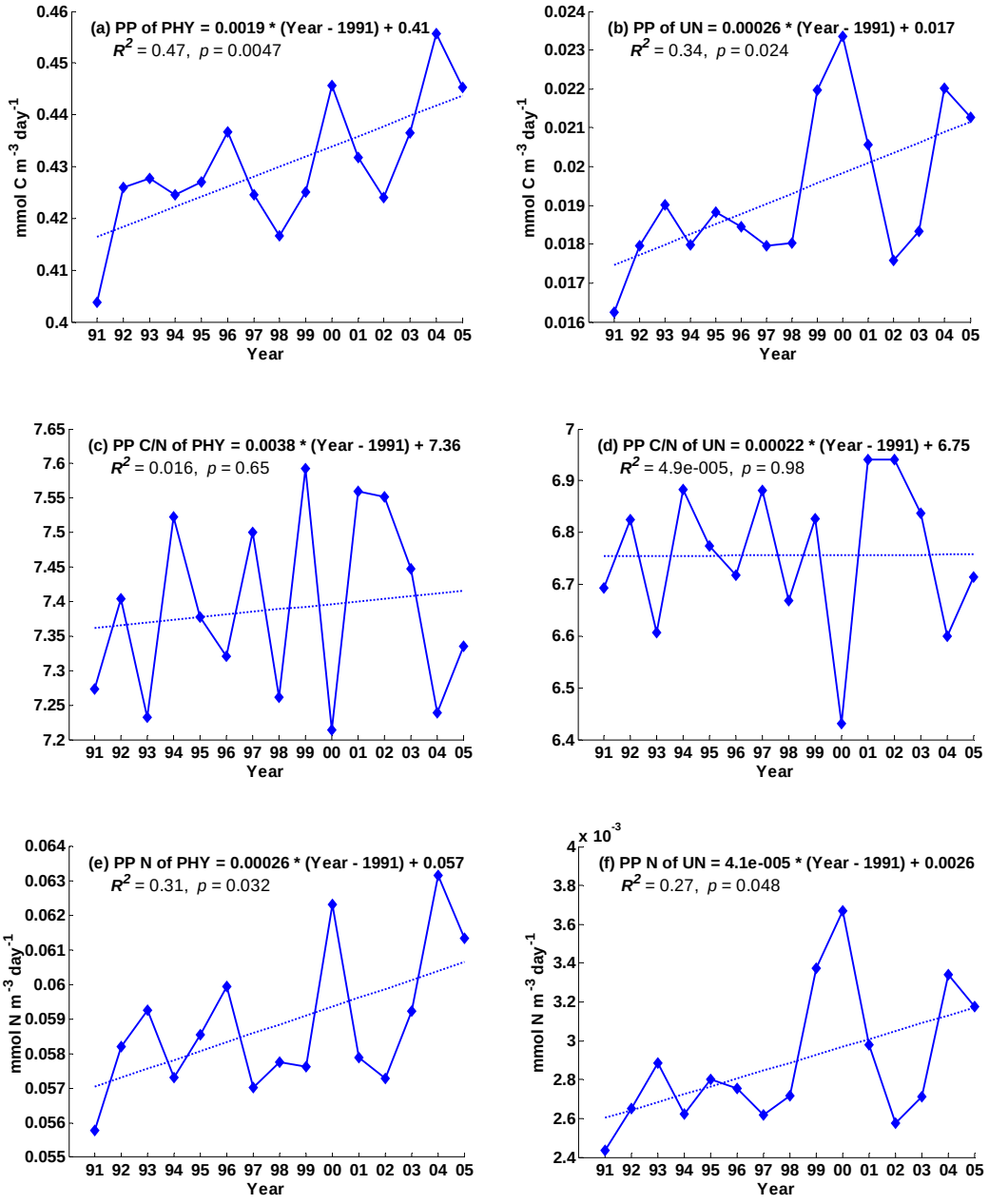

(a) PP of PHY = 0.0019 * (Year - 1991) + 0.41
R2 = 0.47, p = 0.0047
mmol C m-3 day-1
Year
(b) PP of UN = 0.00026 * (Year - 1991) + 0.017
R2 = 0.34, p = 0.024
mmol C m-3 day-1
Year
(c) PP C/N of PHY = 0.0038 * (Year - 1991) + 7.36
R2 = 0.016, p = 0.65
Year
(d) PP C/N of UN = 0.00022 * (Year - 1991) + 6.75
R2 = 4.9e-005, p = 0.98
Year
(e) PP N of PHY = 0.00026 * (Year - 1991) + 0.057
R2 = 0.31, p = 0.032
mmol N m-3 day-1
Year
x 10-3
(f) PP N of UN = 4.1e-005 * (Year - 1991) + 0.0026
R2 = 0.27, p = 0.048
mmol N m-3 day-1
Year

Figure 10

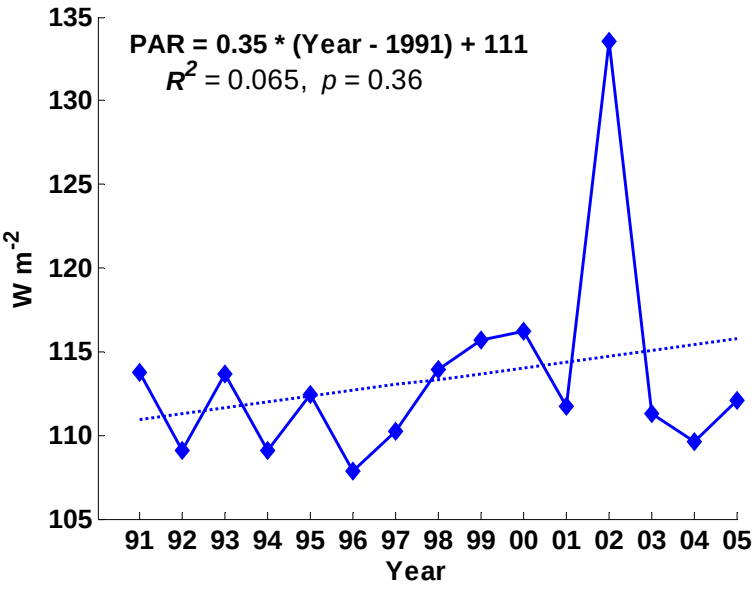

PAR = 0.35 * (Year - 1991) + 111
R2 = 0.065, p = 0.36
W m-2
135
130
125
120
115
110
105
91 92 93 94 95 96 97 98 99 00 01 02 03 04 05
Year

Figure 11

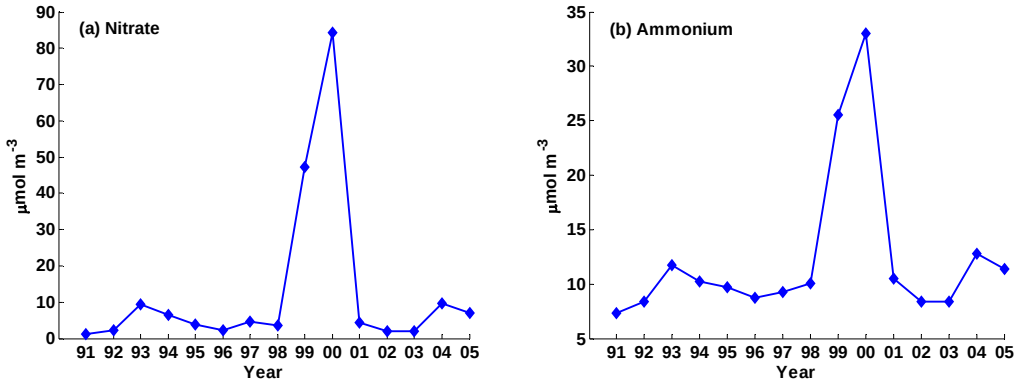

(a) Nitrate
µmol m-3
90
80
70
60
50
40
30
20
10
0
91 92 93 94 95 96 97 98 99 00 01 02 03 04 05
Year
(b) Ammonium
µmol m-3
35
30
25
20
15
10
5
91 92 93 94 95 96 97 98 99 00 01 02 03 04 05
Year

Figure 12

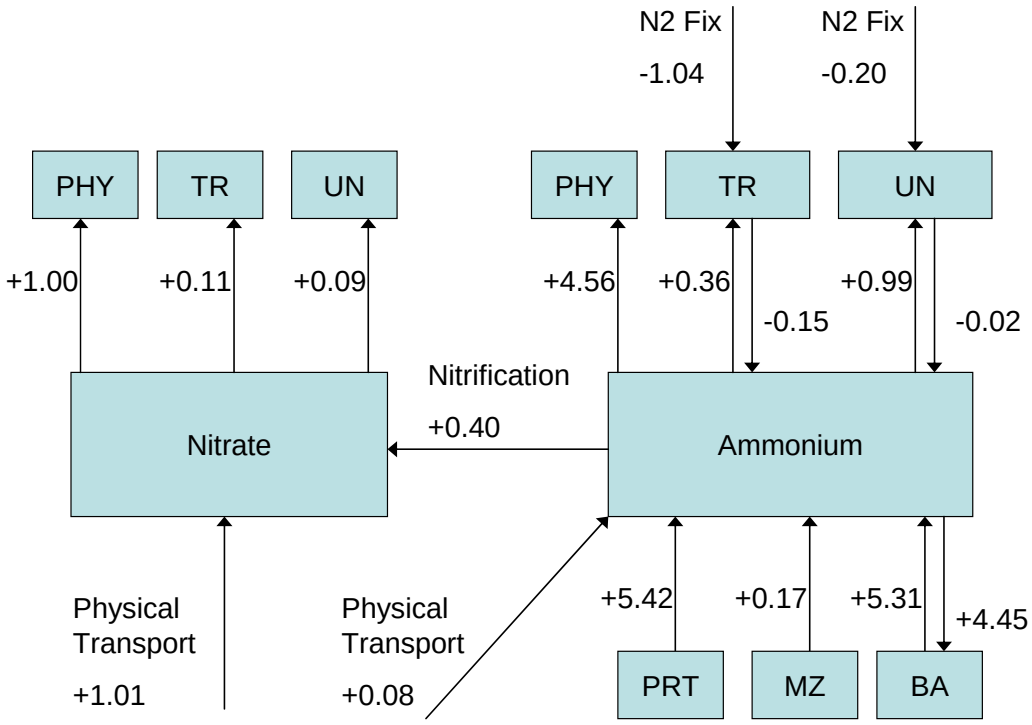

N2 Fix
-1.04
N2 Fix
-0.20
PHY
TR
UN
PHY
TR
UN
+1.00
+0.11
+0.09
+4.56
+0.36
-0.15
+0.99
-0.02
Nitrification
+0.40
Nitrate
Ammonium
Physical
Transport
+1.01
Physical
Transport
+0.08
+5.42
+0.17
+5.31
+4.45
PRT
MZ
BA

Figure 13

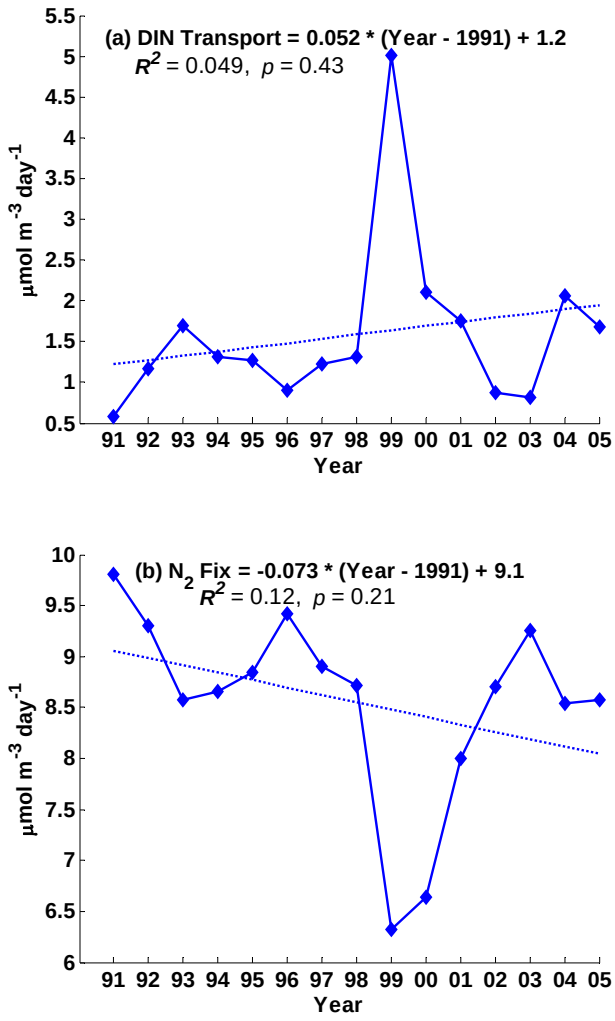

(a) DIN Transport = 0.052 * (Year - 1991) + 1.2
R^2 = 0.049, p = 0.43
µmol m^{-3} day^{-1}
Year
(b) N_2 Fix = -0.073 * (Year - 1991) + 9.1
R^2 = 0.12, p = 0.21
µmol m^{-3} day^{-1}
Year

Figure 14

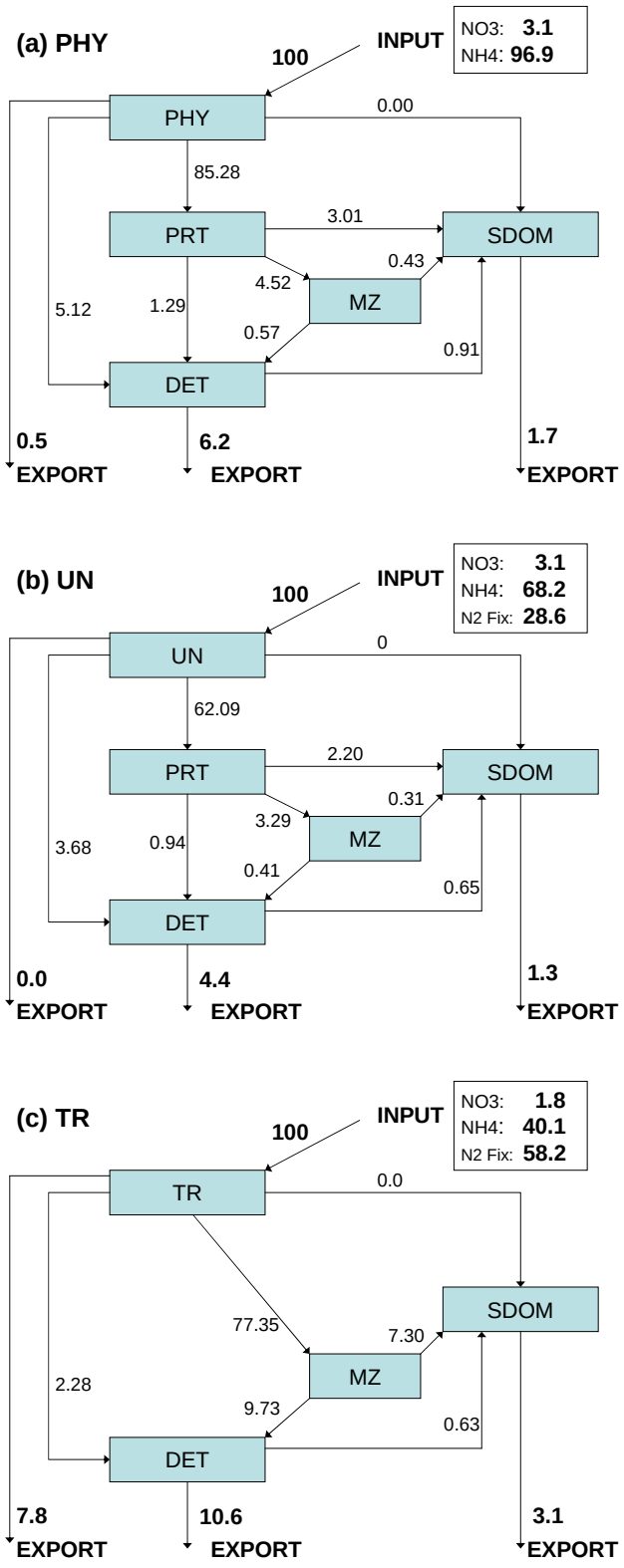

(a) PHY
100
INPUT
NO3: 3.1
NH4: 96.9
PHY
0.00
85.28
PRT
3.01
SDOM
0.43
4.52
MZ
5.12
1.29
0.57
0.91
DET
0.5
EXPORT
6.2
EXPORT
1.7
EXPORT
(b) UN
100
INPUT
NO3: 3.1
NH4: 68.2
N2 Fix: 28.6
UN
0
62.09
PRT
2.20
SDOM
0.31
3.29
MZ
3.68
0.94
0.41
0.65
DET
0.0
EXPORT
4.4
EXPORT
1.3
EXPORT
(c) TR
100
INPUT
NO3: 1.8
NH4: 40.1
N2 Fix: 58.2
TR
0.0
SDOM
77.35
7.30
MZ
2.28
9.73
0.63
DET
7.8
EXPORT
10.6
EXPORT
3.1
EXPORT

Figure 15

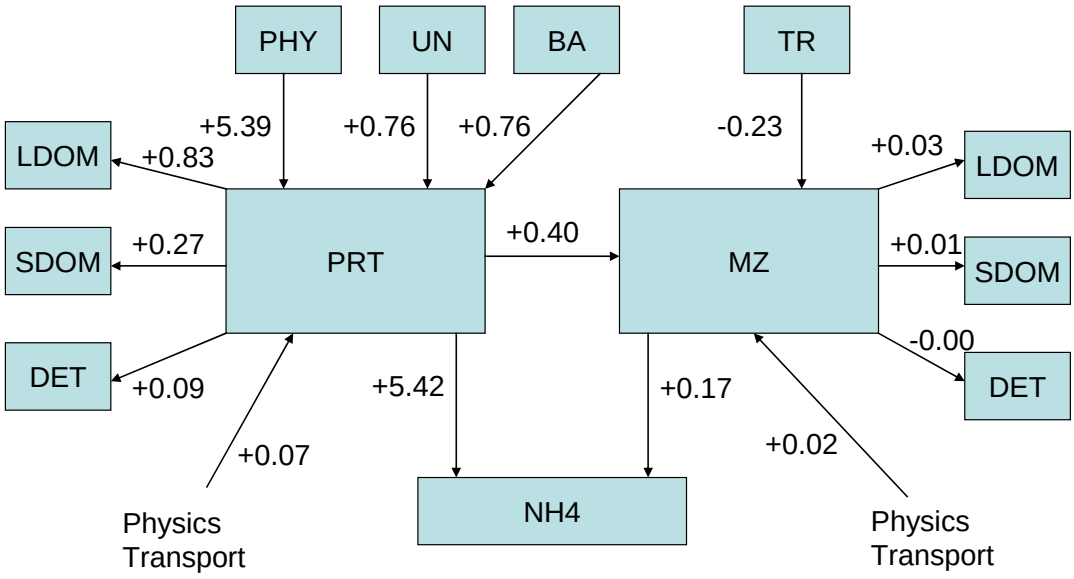

PHY
UN
BA
TR
LDOM
SDOM
DET
PRT
MZ
LDOM
SDOM
DET
NH4
+5.39
+0.76
+0.76
-0.23
+0.83
+0.27
+0.09
+0.07
+0.40
+5.42
+0.17
+0.02
+0.03
+0.01
-0.00
Physics
Transport
Physics
Transport

Figure 16

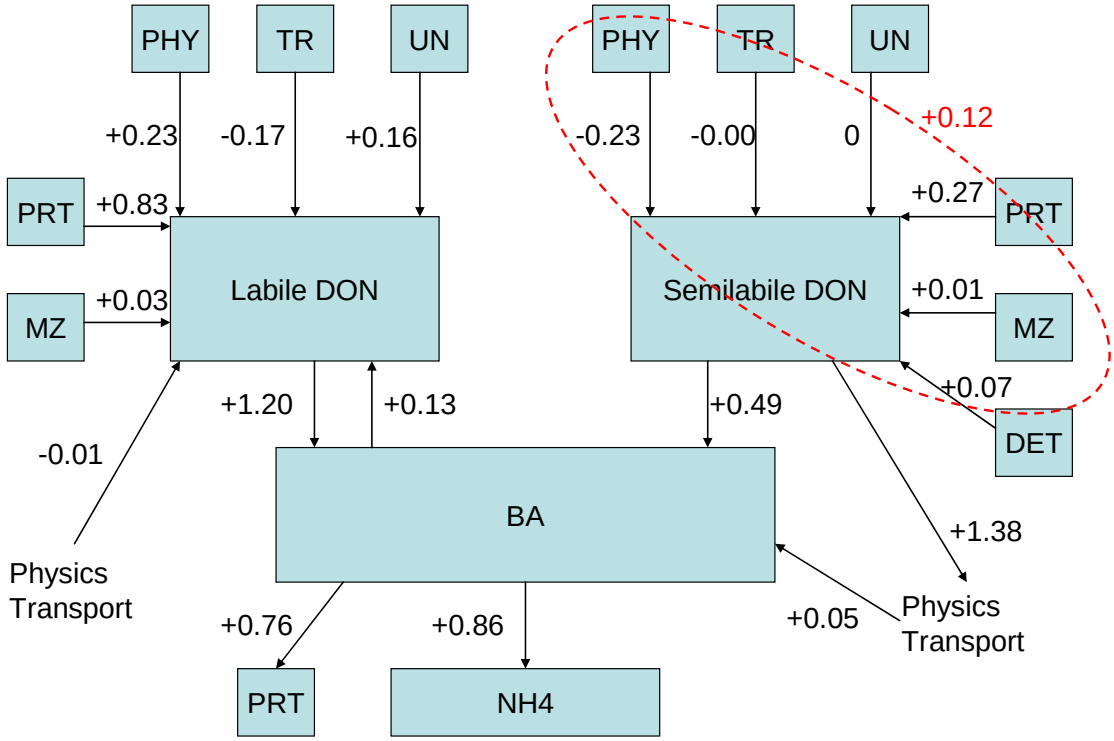

PHY
TR
UN
PHY
TR
UN
+0.23
-0.17
+0.16
-0.23
-0.00
0
+0.12
PRT
+0.83
+0.27
PRT
Labile DON
Semilabile DON
MZ
+0.03
+0.01
MZ
+0.07
DET
+1.20
+0.13
+0.49
-0.01
BA
+1.38
Physics
Transport
+0.76
+0.86
+0.05
Physics
Transport
PRT
NH4

Figure 17

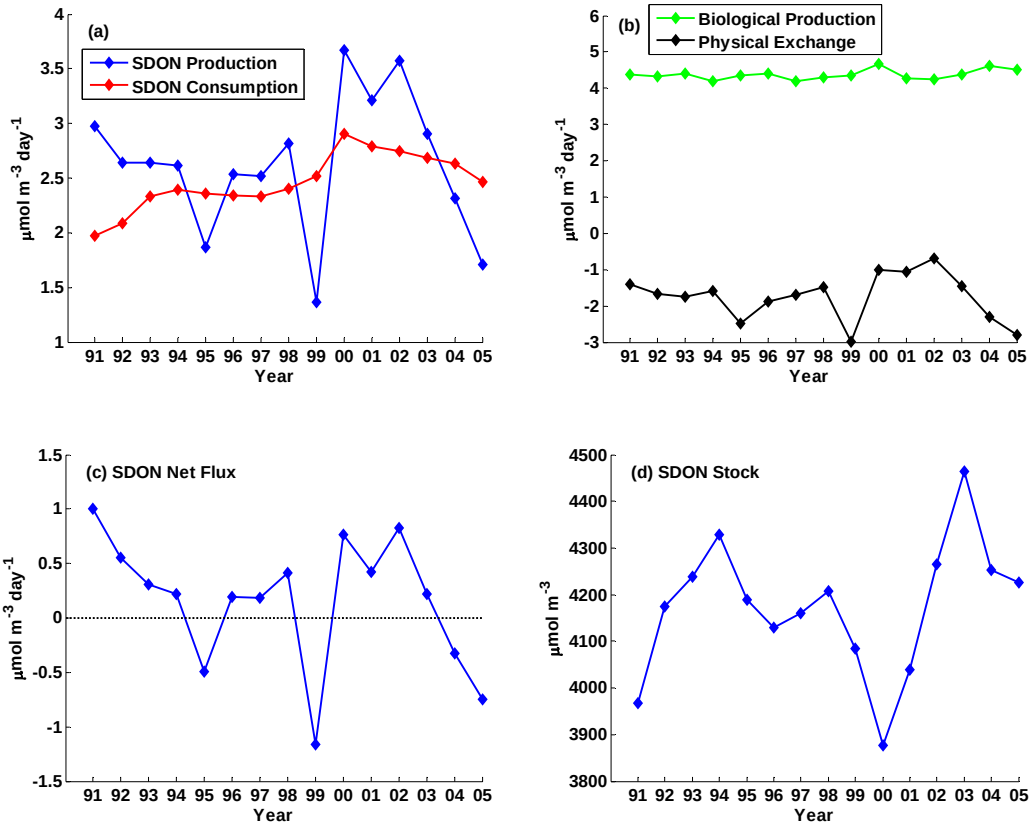

(a)
SDON Production
SDON Consumption
µmol m-3 day-1
Year
(b)
Biological Production
Physical Exchange
µmol m-3 day-1
Year
(c) SDON Net Flux
µmol m-3 day-1
Year
(d) SDON Stock
µmol m-3
Year

Figure 18

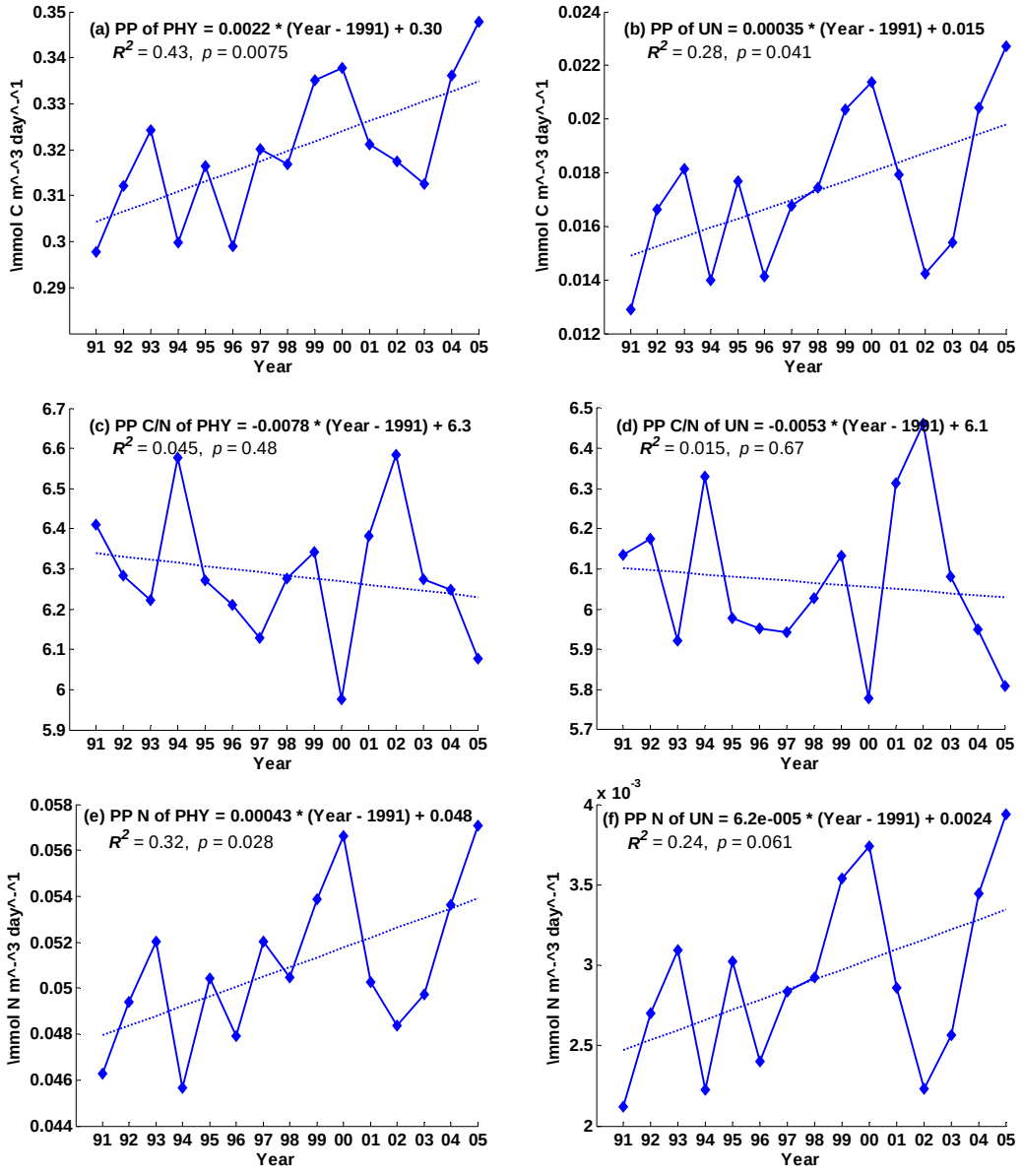

(a) PP of PHY = 0.0022 * (Year - 1991) + 0.30
R2 = 0.43, p = 0.0075
\mmol C m^-^3 day^-^1
Year
(b) PP of UN = 0.00035 * (Year - 1991) + 0.015
R2 = 0.28, p = 0.041
\mmol C m^-^3 day^-^1
Year
(c) PP C/N of PHY = -0.0078 * (Year - 1991) + 6.3
R2 = 0.045, p = 0.48
Year
(d) PP C/N of UN = -0.0053 * (Year - 1991) + 6.1
R2 = 0.015, p = 0.67
Year
(e) PP N of PHY = 0.00043 * (Year - 1991) + 0.048
R2 = 0.32, p = 0.028
\mmol N m^-^3 day^-^1
Year
x 10-3
(f) PP N of UN = 6.2e-005 * (Year - 1991) + 0.0024
R2 = 0.24, p = 0.061
\mmol N m^-^3 day^-^1
Year

Figure 19

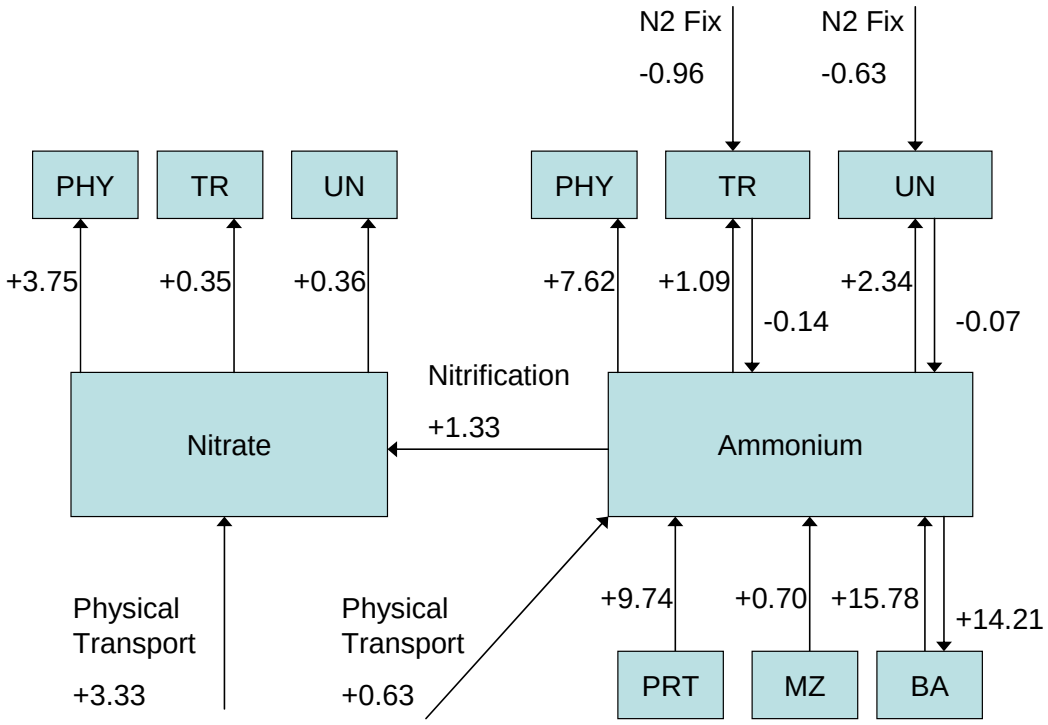

N2 Fix
-0.96
N2 Fix
-0.63
PHY
TR
UN
PHY
TR
UN
+3.75
+0.35
+0.36
+7.62
+1.09
-0.14
+2.34
-0.07
Nitrification
+1.33
Nitrate
Ammonium
Physical
Transport
+3.33
Physical
Transport
+0.63
+9.74
+0.70
+15.78
+14.21
PRT
MZ
BA

Figure 20

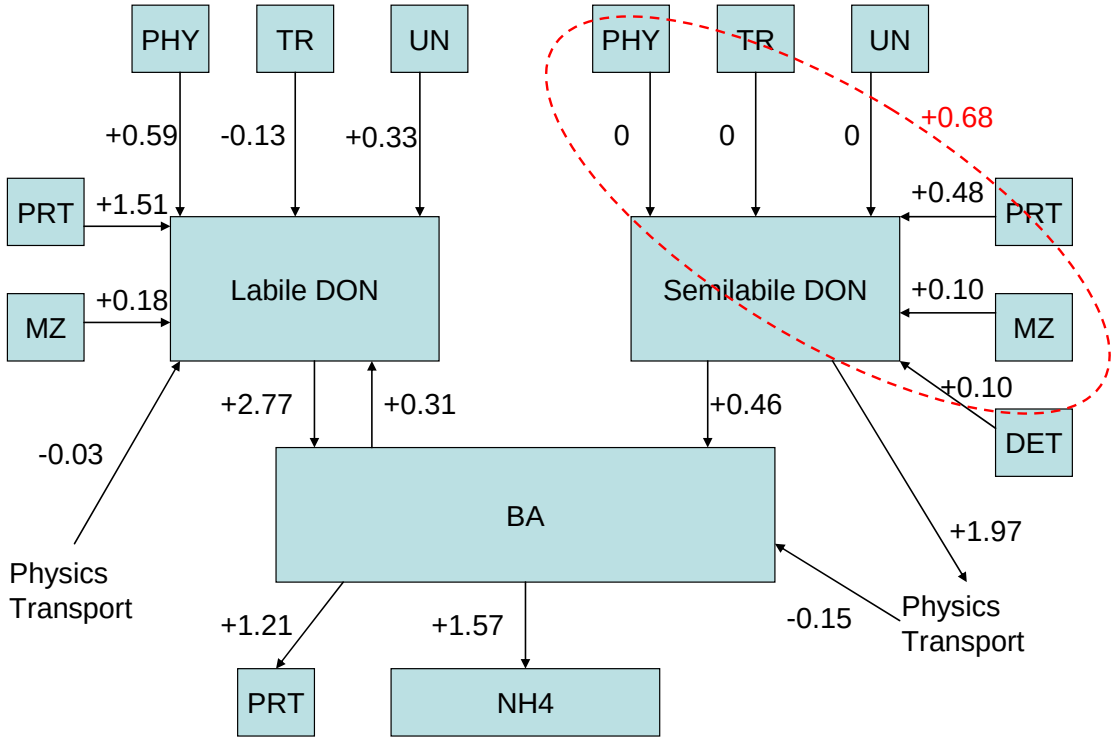
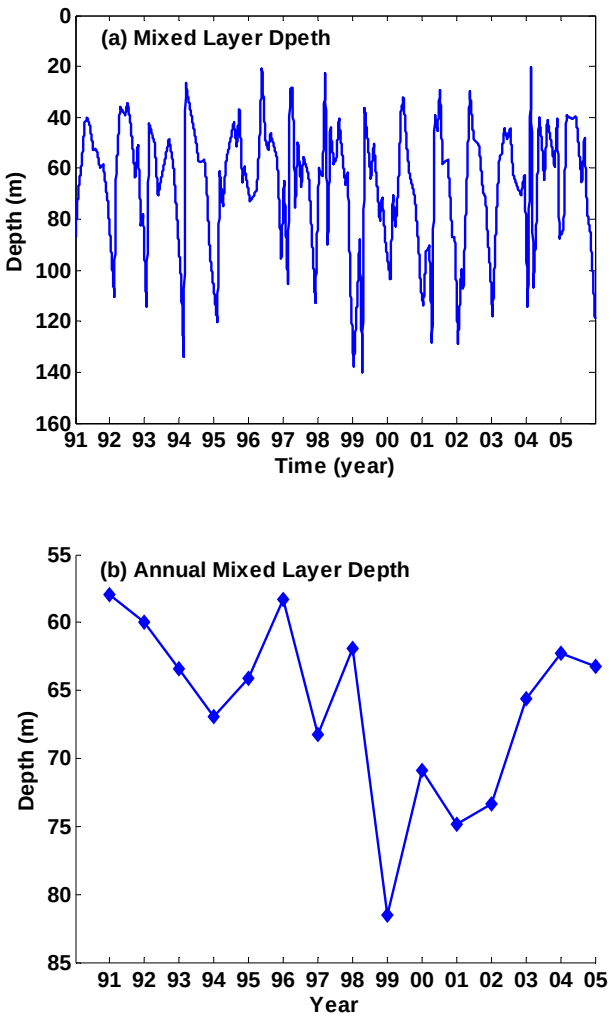

PHY
TR
UN
PHY
TR
UN
+0.59
-0.13
+0.33
0
0
0
+0.68
PRT
+1.51
+0.48
PRT
Labile DON
Semilabile DON
MZ
+0.18
+0.10
MZ
+0.10
DET
+2.77
+0.31
+0.46
-0.03
BA
+1.97
Physics
Transport
+1.21
+1.57
-0.15
Physics
Transport
PRT
NH4

Figure 21


(a) Mixed Layer Dpeth
Depth (m)
0
20
40
60
80
100
120
140
160
91 92 93 94 95 96 97 98 99 00 01 02 03 04 05
Time (year)
(b) Annual Mixed Layer Depth
Depth (m)
55
60
65
70
75
80
85
91 92 93 94 95 96 97 98 99 00 01 02 03 04 05
Year

Chapter 5. Constraining a Marine Ecosystem Model with Heterotrophic Bacterial Production Measurements

Abstract

Heterotrophic bacteria are an essential component of marine ecosystems in carbon cycles and nutrient regeneration. But their production rates are not always routinely measured in the major observatories in the open ocean. To evaluate the importance of such information, an experiment was conducted based on the data assimilating model constructed in Chapter 3, to compare how well a marine ecosystem model can be constrained and optimized when heterotrophic bacterial production (BP) was included in or removed from the assimilated data. The results showed that BP data were useful to strengthen constraints on the model parameters. Without BP data, the model can give incorrect explanations for some mechanisms of ecosystem processes. Currently widely-used radioisotope-based methods require empirical conversion factors to derive BP rates from leucine or thymidine incorporation rates. The conversion factors are difficult to calibrate and can vary in different environments. Therefore, a second experiment was conducted to address this uncertainty by using lower or higher conversion factors to set average observed BP to 5% or 25% of the average observed primary production. The low BP resulted in less fit of the model to other available observations, and the high BP resulted in unrealistic high bacterial growth efficiency and underestimated the bacterial role in regenerating nutrients. The results reiterate the necessity of routine and accurate BP measurements for the ecological studies in the open ocean.

1. Introduction

Heterotrophic *Bacteria* and possibly *Archaea* (termed heterotrophic bacteria hereafter) are the two major consumers of dissolved organic matter (DOM) in marine ecosystems. About 50% of primary production in marine ecosystems (Ducklow & Carlson 1992, Williams 2000) is released as dissolved organic matter (DOM), so it is important to understand the activity of heterotrophic bacteria and their role in the recycling of organic matter. Most of the organic carbon production metabolized by heterotrophic bacteria is respired to CO_2 (mean value 73% ~ 91%; Anderson & Ducklow 2001), and corresponding organic nutrients (e.g., nitrogen and phosphorus) are remineralized to inorganic form to support primary production. Heterotrophic bacteria themselves can be grazed by protists and then by higher trophic levels: a process called the "microbial loop" (Azam et al. 1983).

Net heterotrophic bacterial production (BP) is routinely measured to quantify these processes. BP is usually measured indirectly using radioisotope-labeled precursors of DNA and/or protein synthesis to yield synthesis rates which must be converted to production rates using empirical factors (Ducklow 2000). The two most common approaches use tritium-labeled (^{3}H-) thymidine (Fuhrman & Azam 1982) and ^{3}H-leucine (Kirchman et al. 1985). Empirical conversion factors are applied to convert the isotope incorporation rates (Kirchman & Ducklow 1993, Ducklow 2000). Although conversion factor ideally should be invariant, or else calibrated for each sample with longer-term incubations, literature-based values of the conversion factor tend to be used to simplify experiments and uncertainties are introduced. Actually even the best method for calibrating the conversion factor is still controversial (Kirchman & Ducklow 1993). The

theoretical conversion factor is derived from the total protein content and the ratio of leucine to protein in a "typical" cultured bacterial cell (Simon & Azam 1989). In my previous data assimilating model in Chapter 3, a commonly used "theoretical" conversion factor of 1.5 kg C /mol incorporated leucine (Simon & Azam 1989) was applied to derive BP before the production data were assimilated into the model. This conversion factor may not necessarily be correct for all the three sites. For example, recent work in a temperate coastal ecosystem found the conversion factor of 1.55 kg C /mol leucine would result in a serious overestimation (76%) of annual production rates (Calvo-Díaz & Morán 2009). Data synthesis showed BP in the open ocean ranged in 4 – 26 % of primary production (Ducklow 2000), which could help constrain the uncertainties of BP estimates.

BP is not always measured even at the major open ocean observatories. BP is routinely measured at the Bermuda Atlantic Time-series Study (BATS) by ^{3}H-thymidine incorporation, but is not routinely measured data at the Hawaii Ocean Time-series (HOT). There were only 14 profiles of BP (measured by ^{3}H-leucine) in the 6-year period of 2000 – 2005 at Station ALOHA, while other core parameters have been measured in monthly intervals since 1988. It seems that the importance of BP for marine ecosystem is still not widely recognized in the scientific community.

BP discussed here is a net biomass production estimate without inclusion of respiration. In my ecosystem model constructed in Chapter 2, BP is a derived variable: gross bacterial production minus respiration. Thus it may be expected that a most direct role of BP in model optimization is to constrain the parameter controlling the heterotrophic bacterial respiration rate versus their gross production. But whether BP can

act indirectly to constrain other processes is not clear. Also, with the uncertainties related to BP, it is not clear how the model will perform if the BP were overestimated or underestimated. In this Chapter, I changed the data assimilating model in Chapter 3 to evaluate the model performance by (1) disabling BP assimilation, and (2) increasing and reducing the assimilated BP values. The experiment was conducted to reveal the importance of BP in optimizing marine ecosystem models.

2. Methods

The experiments were applied to all three sites in Chapter 3, the Arabian Sea (AS), Equatorial Pacific (EQP) and Hawaii Ocean Time-series Station ALOHA in year 2002 (HOT). See Chapter 2 for the model structure and data assimilation framework. See Chapter 2 for the model structure and data assimilation framework.

The first experiment was to remove BP from assimilated data list, and then optimize the constrainable model parameters starting from their default values (Table 7 of Chapter 3). Chapter 3 has identified some of model parameters that were (1) optimizable and constrainable, (2) optimizable but not constrainable, or (3) not optimzable. Only the parameters identified as optimizable and constrainable (marked as “CS” in Table 7 of Chapter 3) were optimized in this study and started from their default values. Other parameters were set to their final, optimized if applicable, values as in Table 7, Chapter 3. Thus the purpose of this experiment was to reveal whether the previous constrainable parameters were still constrainable, and whether there was any significant change of the confidence intervals for those constrainable parameters.

As discussed above, BP in the open ocean ranges from 4 – 26% of primary

production (PP). The second experiment was to change the assimilated BP values to match the high/low ends of this observed range. For each site, all the observed BP profiles were multiplied by a constant factor to increase (decrease) values so that the average of the depth-integrated BP would equal to 25% (5%) of the average of the depth-integrated PP observations. The weight in the cost function for the BP was also changed accordingly because this adjustment also changed the mean of the observed BP (see Section 3.2, Chapter 2).Then the experiment re-optimized the model parameters starting from their previous optimized values as in Chapter 3. At first, only those optimizable and constrainable parameters identified in Chapter 3 were optimized. But if after optimization some unoptimized parameters had high gradients for the cost function, these parameters were added to the optimizing list and the model was optimized again. The purpose of this experiment was to reveal whether and how the uncertainty of BP would significantly influence the model results.

3. Results and Discussion

3.1. Experiment 1: No BP assimilated

After removing BP from the assimilated data list, the optimized parameter values and their confidence intervals showed different patterns from the results when BP was assimilated (**Figure 1**). Each parameter value and its confidence interval in Figure 1 was normalized to its parameter value with BP data assimilation. Thus Figure 1 presents the relative change of optimized parameter values after BP was removed from the assimilation, and also the constraint strength for each parameter.

At site AS (**Figure 1a**), after the removal of BP from assimilation, all the

parameters were still optimized to values close to those with BP assimilation. 12 of the 14 parameters changed less than 5%, and other 2 parameter (#10 and #13) changed less than 8%. The confidence intervals for most parameters did not change significantly, except $r^{A}_{\max,BA}$ (Parameter #7), the maximum active respiration rate of heterotrophic bacteria, and g_{BA} (Parameter #10), the half-saturation density of heterotrophic bacteria for protozoan grazing. Without BP assimilation, these two parameters became less constrained. $r^{A}_{\max,BA}$ directly controls the heterotrophic bacterial respiration, and g_{BA} directly controls the removal of heterotrophic bacteria by protozoa. Thus these two parameters are most important to control the BP and heterotrophic bacterial biomass. Although the heterotrophic bacterial biomass was assimilated, the uncertainty of BP generates uncertainty in the grazing pressure on heterotrophic bacteria while fitting the biomass data. With the new values of these parameters, the cost function values for AS did not show significant change (**Table 1**). That is, without assimilating BP, values of modeled state variables in similar to the observations can still be reached, even for BP itself. Thus the removal of BP from assimilation at AS only resulted in two less well-constrained parameters, but did not influence the model results significantly.

At site EQP (**Figure 1b**), after the removal of BP from assimilation, some parameters were optimized to similar values as those with BP assimilation, but some others (Parameters #5 - #9) were optimized to values far from those with BP assimilation. The parameter controlling heterotrophic bacterial respiration, $r^{A}_{\max,BA}$ (Parameter #6), was well constrained with BP assimilated, but it decreased ~ 70% and became much less well-constrained without BP assimilated. With less respiration, the heterotrophic bacterial growth rate was reduced to fit the assimilated heterotrophic bacterial biomass, as

indicated by the lower parameter μ_{BA} (Parameter #5). This changed the balance of growth and removal of heterotrophic bacteria and resulted in higher grazing pressure on them, as indicated by the lower parameter g_{BA} (Parameter # 9). As protozoa grazed on both heterotrophic bacteria and phytoplankton, the lower g_{BA} also changed the parameters μ_{PRT} and g_{PHY} (Parameters #7 & 8), which controlled protozoan growth rate and their grazing pressure on phytoplankton. With these new optimized parameter values, the total cost function for EQP increased (**Table 1**), mostly from the increased cost of BP. As BP was not assimilated in the experiment, the new optimization process did not force the model to fit BP. If the cost of BP was not included, the total cost became 91.83 versus 88.78 with/without BP assimilated. In other words, the model successfully optimized all the assimilated data types, except BP, to similar fit levels as the results with BP assimilated. But this optimization resulted in incorrect BP values. The comparison of average carbon fluxes of the model results with and without BP assimilation (**Figure 2 a & b**) showed that without BP assimilated, heterotrophic bacteria respired significantly less carbon and the heterotrophic bacterial growth efficiency (BGE) was ~ 70%, and protozoa grazed on heterotrophic bacteria much more heavily. The nitrogen fluxes (**Figure 2 c & d**) showed that without BP assimilation heterotrophic bacteria had net uptake of dissolved inorganic nitrogen (DIN), while they had net production of DIN with BP assimilation. The total production of DIN was not significantly influenced, as the organic matter not respired by heterotrophic bacteria was ingested and respired by protozoa. In summary, the removal of BP from assimilation at EQP resulted in different optimized parameter values and the modeled results related to heterotrophic bacteria. The new results were not consistent with the general feature of the open ocean that

heterotrophic bacteria lost their role in remineralizing DIN.

At site HOT, after the removal of BP from assimilation, the parameter controlling the maximum active respiration rate of heterotrophic bacteria, $r^{A}_{\max,BA}$, was optimized to a unrealistic value (indicating the maximum respiration would be only 1% of growth rate), and the confidence interval was $\pm$1,100% of the optimized value. Thus the optimization for this parameter failed. It had to be removed from the optimization and its value was set to the "known" optimized value with BP assimilation. The optimization without BP assimilation for other parameters resulted in almost identical values and confidence intervals as those with BP assimilation (**Figure 1c**). Thus it indicated that BP assimilation only influenced the optimization for $r^{A}_{\max,BA}$ at HOT. If the correct value of this parameter was known, the BP probably would have been unnecessary in order to optimize the model.

3.2. Experiment 2: High and Low BP assimilated

(a) Parameter optimization and fit to observations

When the default conversion factor for BP, 1.5 kgC /mol leucine, was used, the average depth-integrated observed BP was 15.4%, 9.4% and 17.4% of the average observed PP at AS, EQP and HOT respectively (Table 3). As discussed above, the conversion factor was decreased and increased so that this ratio of BP to PP was changed to 5% and 25% at each site. These two types of experiments will be referred as "low BP" and "high BP" hereafter, and the BP derived from the default conversion factor will be referred as "default BP."

Default BP was replaced from low BP and high BP, and the model was optimized again starting from the previously optimized parameter (using default BP).

At AS (**Figure 3a**), as expected, low BP resulted in a higher parameter $r^{A}_{\max,BA}$ (Parameter #7) thus a higher respiration rate, and high BP resulted in a lower $r^{A}_{\max,BA}$ and a lower respiration rate. Low BP also resulted in a lower growth rate of heterotrophic bacteria (lower Parameter #6) and a lower grazing pressure on heterotrophic bacteria (higher Parameter #10) in order to fit the heterotrophic bacterial biomass. The grazing pressure on protozoa (Parameters #12 & #13) also changed with low BP in response to the changed foodweb. These parameters did not change significantly with high BP, except for higher grazing pressure on heterotrophic bacteria (lower Parameter #10). The cost of BP increased 16% with low BP assimilated and did not change with high BP assimilated (**Figure 4a**), which indicated it was more difficult for the model to fit low BP than to fit default and high BP for the AS site. Low BP also resulted in lower fit (higher costs) in nitrate, metazoan zooplankton, PP and semilabile dissolved organic nitrogen (DON). No significant cost change was observed for the high BP. The total cost for low BP increased 8%, while it increased < 1% for high BP.

At EQP (**Figure 3b**), low BP did not change most parameters including heterotrophic bacterial respiration rate, except for a lower grazing pressure on heterotrophic bacteria (higher Parameter #9). But the confidence interval for this parameter was wide and overlapped the value from default BP. High BP resulted more changed parameters including lower growth rate and respiration rate of heterotrophic bacteria (Parameters #5 & #6), higher grazing pressure on heterotrophic bacteria (Parameter #9) and changed grazing related to protozoa and phytoplankton (Parameter #7 and #8). High BP also significantly narrowed the confidence intervals of the parameters related to heterotrophic bacteria growth rate and grazing pressure (Parameter #5 & #9).

The model was easier to fit to high BP (reduced cost) and more difficult to fit low BP (increased cost) (**Figure 4b**). High BP also significantly reduced the costs of, and resulted in better fit to, nitrate, Chlorophyll *a* and heterotrophic bacterial biomass, and resulted in less fit to semilabile dissolved organic carbon (DOC). Low BP resulted in less fit to heterotrophic bacterial biomass, and slightly better fit to semilabile DOC. The total cost for high BP decreased 3% and revealed a better fit, while the total cost for low BP increased 1%.

At HOT (**Figure 3c**), similar to AS, low BP resulted in a higher parameter $r^{A}_{\max,BA}$ (Parameter #7) thus higher respiration rate, and high BP resulted in a lower $r^{A}_{\max,BA}$. High BP resulted in a higher grazing pressure on heterotrophic bacteria (lower Parameter #9) in order to fit the heterotrophic bacterial biomass. When low BP was assimilated, the parameters controlling phytoplankton nitrogen assimilation rate (Parameter #3) and grazing pressure on heterotrophic bacteria (Parameter #9) became not constrainable, and they were set to their values when default BP was assimilated. It was very difficult for the model to fit low BP as indicated by the 58% increase of the cost from BP, while it fitted better to high BP (**Figure 4c**). Low BP also resulted in significantly less fit to nitrate, phosphate, semilabile DOC and DON and particulate phosphorus flux, but had a better fit to particulate organic phosphorus (POP). High BP also resulted in better fit to nitrate, metazoan zooplankton and semilabile dissolved organic phosphorus (DOP), but less fit to semilabile DON. The total cost for low BP increased 7% and revealed a less fit, while the total cost for high BP decreased 1%.

(b) Key model results

In **Table 4**, some important variables of the optimized model results are compared

for assimilating low/high BP versus default BP. In most cases, the model still can maintain the PP, although the PP drops ~ 5% when low BP was assimilated at AS. The resulted BP/PP ratios changed accordingly. Note the realized BP/PP ratios were not 5% and 25% as proposed. The reasons are complicated. First, the low/high BP were constructed based on the observations, which are sparse, and probably not representative of the whole annual cycle; while the results in **Table 4** were the average of the modeled results over the whole model domain. Second, the assimilated PP observations are for particulate PP only, and the model was optimized to fit only its particulate PP to the observed PP. The ratios of modeled BP to modeled particulate PP became closer to the proposed target of 5% and 25% at EQP and HOT. But at AS, the ratios using particulate PP became further away from the proposed target, which indicated less fidelity of the modeled BP and PP to observations at AS.

The most important potential ways to fit the low/high BP in the model could be realized either by changing heterotrophic bacterial growth efficiency (BGE), or by changing the resource consumption rate (termed the bacterial carbon demand, BCD), or both. The changes were generated by the optimization automatically and objectively. As expected, low BGE was realized for low BP and high BGE for high BP (**Table 4**). The BGE for high BP reached ~ 50% at all the three sites, which may be too high for the open ocean (del Giorgio & Cole 1998), while the BGE values for low BP seems more reasonable. Different BGE resulted in different carbon respiration rates by heterotrophic bacteria. Thus different BP values will influence the conclusions about the role of heterotrophic bacteria in the carbon cycle. High BP also resulted in higher rates of labile DOC-supported BCD (L-BCD), and lower BP resulted in lower L-BCD. Because of the

short residence time of labile DOC and therefore the tight coupling of its consumption and production, the higher L-BCD indicated a higher production of labile DOC and an intensified recycling, mostly via the heterotrophic bacteria – protozoa – labile DOM pathway. The same pattern also applied to semilabile DOC. Thus the model was able to adjust BP partly from controlling the recycling through this pathway without significantly influencing other components of the model.

The changed BP did not influence the export except for the low BP experiment for AS (**Table 4**), when higher export of dissolved organic matter was generated. That was because heterotrophic bacteria consumed less semilabile DON and the biological production of semilabile DON significantly exceeded its consumption rate, thus the excess DON was exported.

Another important role of heterotrophic bacteria is to recycle inorganic nutrients and thus support PP (Kirchman 2000, Landry 2001). The rate of heterotrophic bacteria-recycled nitrogen changed with low/high BP values assimilated. Low BP generated low BGE and higher respiration and thus higher recycled nitrogen rates, and vice versa. The rate, in percentage of nitrogen requirement of PP, showed large differences (~ 2 – 4 times) between the experiments with low and high BP (**Table 4**). The total recycled nitrogen did not change (data not shown), because the high BP related to high removal rates by protozoa and high nitrogen recycling by protozoa, and vice versa for low BP. That is, the model would predict a compensating effect of nitrogen recycling by zooplankton if heterotrophic bacteria recycle less. Thus different BP will impact the evaluation of the role of heterotrophic bacteria in nutrient recycling.

3.3. Necessity of BP measurements?

Bacterial ecologists have advocated measuring BP as an essential component of aquatic ecosystems because of its importance for the microbial loop and carbon and other elemental cycling (Cole & Pace 1995, Ducklow 2000), as also suggested by the previous chapters of this dissertation. In the present chapter, a first experiment was conducted to quantify how the model optimization would act in the absence of BP. The parameter controlling heterotrophic bacterial respiration rate (and therefore BGE) was affected most. After a certain amount of DOM is produced for heterotrophic bacterial usage, the model without BP information did not have direct constraints on how much of the consumed carbon and other elements should be respired. At all the three study sites, this parameter had higher uncertainties, and was even completely unconstrainable at HOT (**Figure 1**). The results indicated different implications of BP in terms of constraining the model. The other observations without BP were good enough at AS to constrain the model to a correct condition, while they generated incorrect patterns at other two sites. The experiments in the absence of BP in all the three sites could still obtain the same level of fit to other observations (**Table 2**). That is, even if the model can successfully fit other observations, the internal mechanisms, especially those related to heterotrophic bacterial activities, could be incorrect. For example, the results at EQP would incorrectly indicate heterotrophic bacteria as a net consumer of inorganic nutrients instead of a net producer if BP was not used for constraining the model.

BP in the model was a derived variable that equals to BCD minus BR. Thus the constraint on BP is indirect. Ideally if both BP and heterotrophic bacterial respiration (BR) can both be measured, the model would certainly be more strictly and directly constrained. Without BR measurements, BP alone may not improve understanding of the

role of heterotrophic bacteria in aquatic systems (Jahnke & Craven 1995). Unfortunately, it is easy to measure BP, but difficult to routinely measure BR. The size of heterotrophic bacteria is ~ 0.6 μm in diameter which is close to some autotrophs and protists (Gasol et al. 2008). So accurate isolation of the heterotrophic bacteria is difficult and there is no unambiguous selection of filter pore size (ranges from 0.6 to 3μm) to separate algae, grazers and bacteria (Robinson 2008). In addition, experimental isolation of the heterotrophic bacteria breaks the trophic linkage to predators and organic and inorganic nutrient suppliers (Robinson 2008). Nonetheless, this study showed that BP certainly provided constrains on the model. Even without respiration measurements, other observations in different aspects of the ecosystem indirectly constrained the production and thus consumption of dissolved organic matter by heterotrophic bacteria, i.e., BCD. Certainly BP + BR is a better and more strict constraint on BCD. But with current methodological limitations on BR measurements, it has to be emphasized that routine BP measurement is necessary to improve marine ecosystem models.

3.4. Uncertainty of BP

Commonly used empirical and theoretical conversion factors for leucine incorporation to BP range over at least a factor of 4: 0.7 – 3.1 kg C/mol leucine (Simon & Azam 1989, Kirchman & Ducklow 1993, González et al. 2003, Peréz et al. 2005). This study did not use thymidine-derived BP, but the conversion factors for thymidine also range over about the same range as leucine (Robinson 2008). With the theoretical conversion factor of 1.5 kg C/mol leucine, the derived BP at all the three sites was in a reasonable range in terms of BP/PP ratios (**Table 3**). Although BP/PP at EQP seemed significantly lower than the other two sites, the uncertainty of the conversion factor could

blur the difference. Although BP/PP was highest in HOT, there is some evidence indicating that the conversion factor could be much lower in low productivity, oligotrophic habitats (Pedrós-Alió et al. 1999, Alonso-Saez et al. 2007). Thus it was possible that the BP at HOT was overestimated by using the theoretical conversion factor, as its PP was much lower than other two sites. In addition, seasonal variability may further complicate the determination of correct conversion factor (Calvo-Díaz & Morán 2009).

In this study, the way to address the uncertainty of the conversion factor, and thus BP, was to evaluate whether the lower BP and higher BP values could possibly fit into the modeled ecosystem without impacting the already-realized fit to other observations and without generating unrealistic model results. The new conversion factors in order to set up low/high BP experiments ranged 0.4 – 4.0 kg C/mol leucine (**Table 3**), which were not significantly out of the normally-used range. The low BP experiments at AS and HOT resulted in less fit to other observations, while the high BP worked reasonably but did not clearly improve the fit. Thus the low BP (5% of PP) was not able to fit the current model structure and observations at AS and HOT, while the values around the default BP or higher did not impact the internal consistency of the model. At EQP, the high BP experiments resulted in improvement of the fit to other observations. Thus the default BP limited the model performance and the model preferred higher BP than the default BP. It is likely true because EQP has the lowest BP/PP ratio in all the three sites when default BP was used (**Table 3**). However, the high BP experiments all resulted in high BGE of about 50% (**Table 4**), higher than the normally observed values in the open ocean (del Giorgio & Cole 1998). Thus the model suggests that the high BP/PP of 25% is not

acceptable. Earlier analyses also suggested that in the open sea, in the absence of external input of organic matter, BP/PP is normally below 20% (Anderson & Ducklow 2001). Thus even if high BP improves the fit to observations, it could result from an incorrect mechanism. Different BP rates can change the evaluation of nutrient recycling rates by heterotrophic bacteria (**Table 4**). Higher BP could also lead to conclude that heterotrophic bacteria are less important in recycling nutrients and that zooplankton are more important nutrient recyclers, and vice versa for lower BP (**Table 4**). Thus reliable BP is necessary to correctly assess this issue.

Here the experiments suggested that both the low and high end of BP/PP (5% and 25%) and the corresponding conversion factors of 0.4 – 4.0 kgC /mol leucine reduced the model performance in either less fit to other observations or incorrect mechanisms. The experiments gave a rough constrain and supported the results from previous studies that BP/PP in the open ocean was in 5% - 25% and the commonly used conversion factor of 0.7 – 3.1 kg C/mol leucine was reasonable.

4. Conclusion

Although heterotrophic bacteria are an essential component of marine ecosystems, particularly for carbon cycling and nutrient regeneration, the benefits of routinely measuring their net production are still unclear. This study showed that measurements of heterotrophic bacterial net production (BP), even without measurements of their respiration rates, can help to constrain model parameters and optimize the model results to a correct status. Current widely-used practice requires conversion factors to derive heterotrophic bacterial production from the incorporation rates of isotope-labeled leucine or thymidine, but the conversion factors may vary in different conditions. This study

suggested that the commonly used “theoretical” conversion factor of 1.5 kg C/mol leucine may not necessarily be correct for different sites and can potentially limit model performance. Meanwhile, the high end of the normally-used conversion factor range and the corresponding 25% of BP/PP resulted in unrealistically high heterotrophic bacterial growth efficiency and underestimate the bacterial role in recycling nutrients. The low end of the conversion factor and the corresponding 5% of BP/PP resulted in worse performance of the model and less fit to other observations, which indicates that low BP cannot fit into our current knowledge of the marine ecosystems. Thus although it is certainly better if the conversion factor can be calibrated for each study, the real conversion factor should not be greatly different from the “theoretical” value of 1.5 kg C/mol leucine. The BP/PP ratio probably should be in 5 – 25% in most open ocean environments.

References

Alonso-Saez L, Gasol JM, Aristegui J, Vilas JC, Vaque D, Duarte CM, Agusti S (2007) Large-scale variability in surface bacterial carbon demand and growth efficiency in the subtropical northeast Atlantic Ocean. Limnol. Oceanogr. 52:533-546

Anderson TR, Ducklow HW (2001) Microbial loop carbon cycling in ocean environments studied using a simple steady state model. Aquat. Microb. Ecol. 26:37-49

Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad F (1983) The ecological role of water-column microbes in the sea. Mar. Ecol. Prog. Ser. 10:257-263

Calvo-Díaz A, Morán XAG (2009) Empirical leucine-to-carbon conversion factors for estimating bacterial heterotrophic production: seasonality and predictability in a temperate coastal ecosystem. Appl. Envir. Microbiol.:doi:10.1128/AEM.01570-08

Cole JJ, Pace ML (1995) Why measure bacterial production? A reply to the comment by Jahnke and Craven. Limnol. Oceanogr. 40:441-444

del Giorgio PA, Cole JJ (1998) Bacterial Growth Efficiency in Natural Aquatic Systems. Ann. Rev. Ecol. Syst. 29:503-541

Ducklow HW (2000) Bacterioplankton production and biomass in the oceans. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley, New York, p 85-120

Ducklow HW, Carlson CA (1992) Oceanic bacterial productivity. Adv. Microb. Ecol. 12:113-181

Fuhrman JA, Azam F (1982) Thymidine incorporation as a measure of heterotrophic

bacterioplankton production in marine surface waters: evaluation and field results. Mar. Biol. 66:109-120

Gasol JM, Pinhassi J, Alonso-Sáez L, Ducklow H, Herndl GJ, Koblízek M, Labrenz M, Luo Y, Morán XAG, Reinthaler T, Simon M (2008) Towards a better understanding of microbial carbon flux in the sea. Aquat. Microb. Ecol. 53:21-38

González N, Anadón R, Viesca L (2003) Carbon flux through the microbial community in a temperate sea during summer: role of bacterial metabolism. Aquat. Microb. Ecol. 33:117-126

Jahnke RA, Craven DB (1995) Quantifying the role of heterotrophic bacteria in the carbon cycle: A need for respiration measurements. Limnol. Oceanogr. 40:436-441

Kirchman D, K'Nees E, Hodson R (1985) Leucine incorporation and its potential as a measure of protein synthesis by bacteria in natural aquatic systems. Appl. Envir. Microbiol. 49:599-607

Kirchman DL (2000) Uptake and regeneration of inorganic nutrients by marine heterotrophic bacteria. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, Inc., New York, p 261-288

Kirchman DL, Ducklow HW (1993) Estimating conversion factors for the Thymidine and Leucine methods for measuring bacterial production. In: Kemp PF, Sherr BF, Sherr EB, Cole JJ (eds) Handbook of Methods in Aquatic Microbial Ecology. Lewis Publishers, Boca Raton, FL, p 513-518

Landry M (2001) Microbial loops. In: Steele JH, Thrope SA, Turekian KK (eds) Encyclopedia of Ocean Sciences. Academic Press, London, p 1763-1770

Pedrós-Alió C, Calderón-Paz J-I, Guixa-Boixereu N, Estrada M, Gasol JM (1999) Bacterioplankton and phytoplankton biomass and production during summer stratification in the northwestern Mediterranean Sea. Deep-Sea Res. I 46:985-1019

Peréz V, Fernández E, Marañón E, Serret P, Varela R, Bode A, Varela M, Varela MM, Morán XAG, Woodward EMS, Kitidis V, Garcia-Soto C (2005) Latitudinal distribution of microbial plankton abundance, production, and respiration in the Equatorial Atlantic in Autumn 2000. Deep-Sea Res. I 52:861-880

Robinson C (2008) Heterotrophic bacterial respiration. In: Kirchman DL (ed) Microbial Ecology of the Oceans (second edition). John Wiley & Sons, Inc., Hoboken, New Jersey, p 299 - 334

Simon M, Azam F (1989) Protein content and protein synthesis rates of planktonic marine bacteria. Mar. Ecol. Prog. Ser. 51:201-213

Williams PJl (2000) Heterotrophic bacteria and the dynamics of dissolved organic material. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, New York, p 153- 200

Table 1. List of the parameter names of Figure 1 for sites AS, EQP and HOT. Please refer to Table 2, Chapter 2 for the descriptions of the parameters.

Parameter #	AS	EQP	HOT
1	μ_{PHY}	μ_{PHY}	μ_{PHY}
2	α_{PHY}^{CHL}	α_{PHY}^{CHL}	α_{PHY}^{CHL}
3	θ	$v_{REF,PHY}^{N}$	$v_{REF,PHY}^{N}$
4	pom_{PHY}	$b_{LABI,P}$	$v_{REF,PHY}^{P}$
5	$b_{LABI,P}$	μ_{BA}	θ
6	μ_{BA}	$r_{\max,BA}^{A}$	$b_{LABI,P}$
7	$r_{\max,BA}^{A}$	μ_{PRT}	$r_{\max,BA}^{A}$
8	μ_{PRT}	g_{PHY}	g_{PHY}
9	g_{PHY}	g_{BA}	g_{BA}
10	g_{BA}	$resp_{PRT}^{A}$	$resp_{PRT}^{A}$
11	$resp_{PRT}^{A}$	μ_{MZ}	g_{TR}
12	μ_{MZ}	g_{PRT}	$wnsv$
13	g_{PRT}	$resp_{MZ}^{A}$	
14	$diss$	prf_{P}	
15		$diss$	

Table 2. Change of the cost functions from each assimilated data type after heterotrophic bacterial production (BP) was removed from assimilation at the three sites AS, EQP and HOT.

Data Type	AS			EQP			HOT		
	With BP	Without BP	Change	With BP	Without BP	Change	With BP	Without BP	Change
NO3	8.68	8.69	0.1%	5.38	4.07	-24.4%	1.13	0.98	-13.4%
PO4	2.70	2.83	5.0%	4.00	3.81	-4.9%	0.88	1.07	22.1%
MZc	1.10	1.09	-1.4%	5.31	5.07	-4.5%	1.22	1.01	-16.8%
PHYn	37.86	37.79	-0.2%	16.24	16.09	-0.9%	1.05	1.02	-2.6%
CHL	4.23	4.23	-0.1%	11.88	8.16	-31.3%	10.46	10.33	-1.3%
PP	10.91	11.06	1.4%	19.29	19.33	0.2%	23.05	22.79	-1.1%
BAc	5.76	5.58	-3.0%	1.09	0.83	-23.6%	14.58	14.84	1.8%
BP	7.01	7.30	4.1%	3.26	118.18	3525%	2.13	2.28	7.0%
sDOC	4.17	4.10	-1.7%	8.26	10.36	25.4%	1.19	1.27	6.9%
sDON	2.02	2.10	3.9%				2.39	3.35	40.3%
sDOP							1.27	1.17	-7.9%
POC	2.43	2.38	-2.1%	3.40	3.52	3.5%	2.41	2.55	5.7%
PON	4.82	4.82	0.0%	10.92	11.10	1.7%	1.56	1.65	5.5%
POP							2.29	1.88	-17.8%
STc	5.57	5.55	-0.4%	2.76	2.89	4.9%	0.73	0.74	0.9%
STn	4.97	4.96	-0.2%	3.30	3.54	7.4%	0.71	0.71	0.6%
STp							0.42	0.45	7.5%
Total	**102.23**	**102.47**	**0.2%**	**95.09**	**206.96**	**117.6%**	**67.47**	**68.10**	**0.9%**
Total no BP	**95.22**	**95.17**	**-0.1%**	**91.83**	**88.78**	**-3.3%**	**65.34**	**65.82**	**0.7%**

NO3: Nitrate; **PO4**: phosphate; **MZc**: metazoan zooplankton carbon biomass; **PHYn**: phytoplankton nitrogen biomass; **CHL:** Chloropyll *a*; **PP**: primary production; **BAc**: heterotrophic bacterial carbon biomas; **BP**: heterotrophic bacterial production; **sDOC, sDON** & **sDOP**: semilabile dissolved organic carbon, nitrogen and phosphrous; **POC, PON** & **POP**: particulate organic carbon, nitrogen and phosphorus; **STc, STn** & **STp**: particle carbon, nitrogen and phosphorus flux collected by sediment traps.

Table 3. Averages of observed depth-integrated heterotrophic bacterial production (BP) (mg C m^{-2} day^{-1}) using the default conversion factor value of 1.5 kgC /mol leucine, primary production (PP) (mg C m^{-2} day^{-1}) , their ratios and conversion factors (CF, kg C/mol leucine) to set up the experiments of low BP and high BP at sites AS, EQP and HOT.

Site	BP	PP	BP/PP	CF Low-BP	CF High-BP
AS	205	1330	15.4%	0.49	2.44
EQP	114	1217	9.4%	0.80	4.00
HOT	101	582	17.4%	0.43	2.16

Table 4. Major optimized model results when low BP (L-BP), default BP (D-BP) and high BP (H-BP) are assimilated at sites AS, EQP and HOT. Listed variables are: total primary production (PP, including particulate and dissolved), the ratio of BP to total PP and to particulate PP, heterotrophic bacterial growth efficiency (BGE), heterotrophic bacterial carbon demand supported by labile DOC (L-BCD) and semilabile DOC (S-BCD), the export in particulate and dissolved forms as percentage of PP, heterotrophic bacteria-recycled nitrogen (N) in percentage of N requirement by PP.

Property	**AS**			**EQP**			**HOT**		
	L-BP	D-BP	H-BP	L-BP	D-BP	H-BP	L-BP	D-BP	H-BP
PP (mmol C m^{-3} d^{-1})	0.691	0.724	0.724	0.868	0.866	0.863	0.320	0.314	0.314
BP/PP	7.2%	17.1%	26.9%	3.3%	6.7%	18.7%	3.5%	13.0%	18.6%
BP/Particulate PP	9.1%	21.9%	34.3%	3.8%	7.8%	21.8%	4.3%	15.7%	22.5%
BGE	16.6%	33.8%	50.0%	9.9%	19.6%	50.4%	11.0%	35.2%	48.7%
L-BCD (mmol C m^{-3} d^{-1})	0.220	0.248	0.259	0.182	0.186	0.199	0.090	0.103	0.106
S-BCD (mmol C m^{-3} d^{-1})	0.081	0.119	0.131	0.107	0.110	0.121	0.012	0.013	0.013
Particulate Export (% of PP)	15.3%	15.2%	15.1%	18.1%	18.3%	18.7%	4.1%	4.4%	4.4%
Dissolved Export (% of PP)	3.7%	1.2%	1.1%	8.3%	8.2%	7.9%	1.7%	1.5%	1.4%
Bacteria-recycled N (% of PP)	36.4%	30.1%	18.6%	26.6%	21.9%	7.7%	27.1%	20.9%	15.2%

Figure Legends

Figure 1. Optimized parameter values and their confidence intervals when heterotrophic bacterial production data (BP) were assimilated (blue circles) or were not assimilated (red squares) at the three experimental sites AS (upper panel), EQP (middle panel) and HOT (lower panel). All the optimized parameter values and confidence intervals were normalized to the optimized parameter values with BP assimilation (thus all the optimized parameter values with BP assimilation equal 1 in the figure). The dashed circles mark the parameter most directly controlling the respiration of heterotrophic bacteria: the maximum active respiration rate of heterotrophic bacteria. The parameters in this figure are listed in Table 1.

Figure 2. Comparison of vertical and temporal averages of modeled carbon (C) and nitrogen (N) fluxes (arrows) and stocks (boxes) normalized to C and N components of primary production (stocks normalized to amount of primary production during 1 day) at site EQP. With BP assimilation: panel (a) and (c) for C and N; without BP assimilation, panel (b) and (d) for C and N.

Figure 3. Optimized parameter values and their confidence intervals when default heterotrophic bacterial production data (BP) (blue circles), low BP (green downward triangles) and high BP (red upward triangles) were assimilated at AS (upper panel), EQP (middle panel) and HOT (lower panel). All the optimized parameter values and

confidence intervals were normalized to the optimized parameter values with default BP assimilation. The dashed circles mark the parameter of maximum active respiration rate of heterotrophic bacteria. The parameter names in this figure are listed in Table 1.

Figure 4. The relative changes of each cost component and total cost when the assimilated heterotrophic bacterial production was changed from the default level to lower (green) or higher (red) levels (see text for details of the change) at sites AS (a), EQP (b) and HOT (c). Acronyms are same as Table 2.

Figure 1

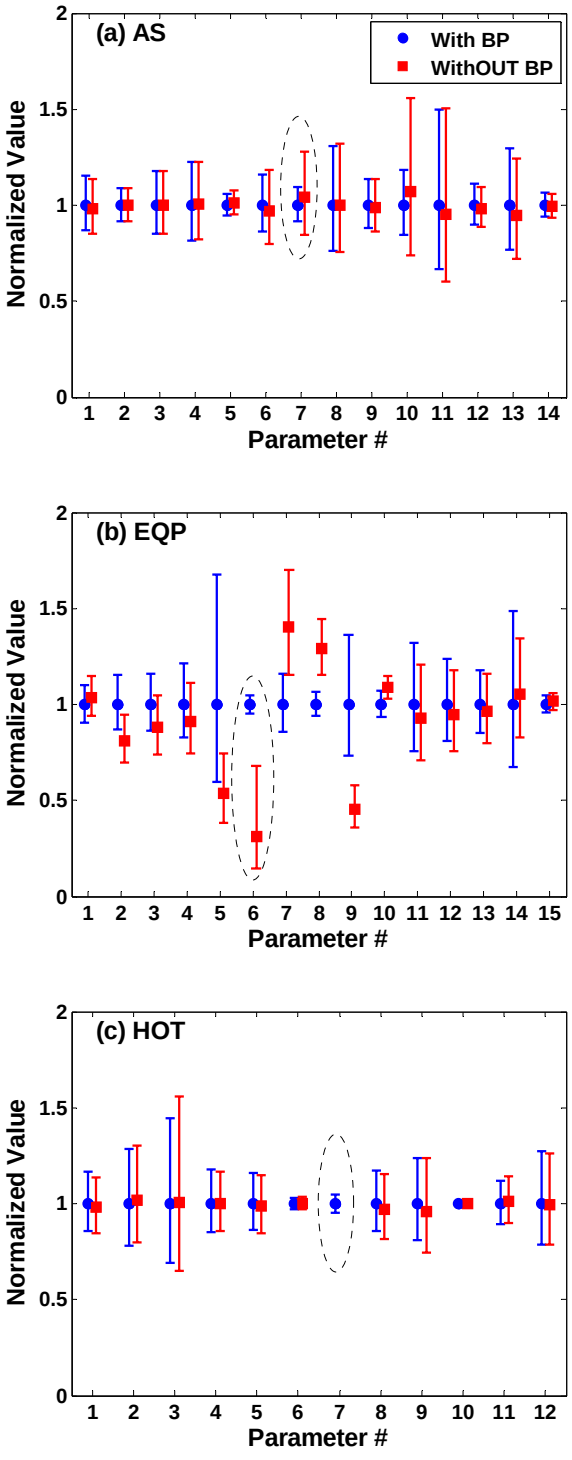

(a) AS
With BP
WithOUT BP
Normalized Value
Parameter #
(b) EQP
Normalized Value
Parameter #
(c) HOT
Normalized Value
Parameter #

Figure 2

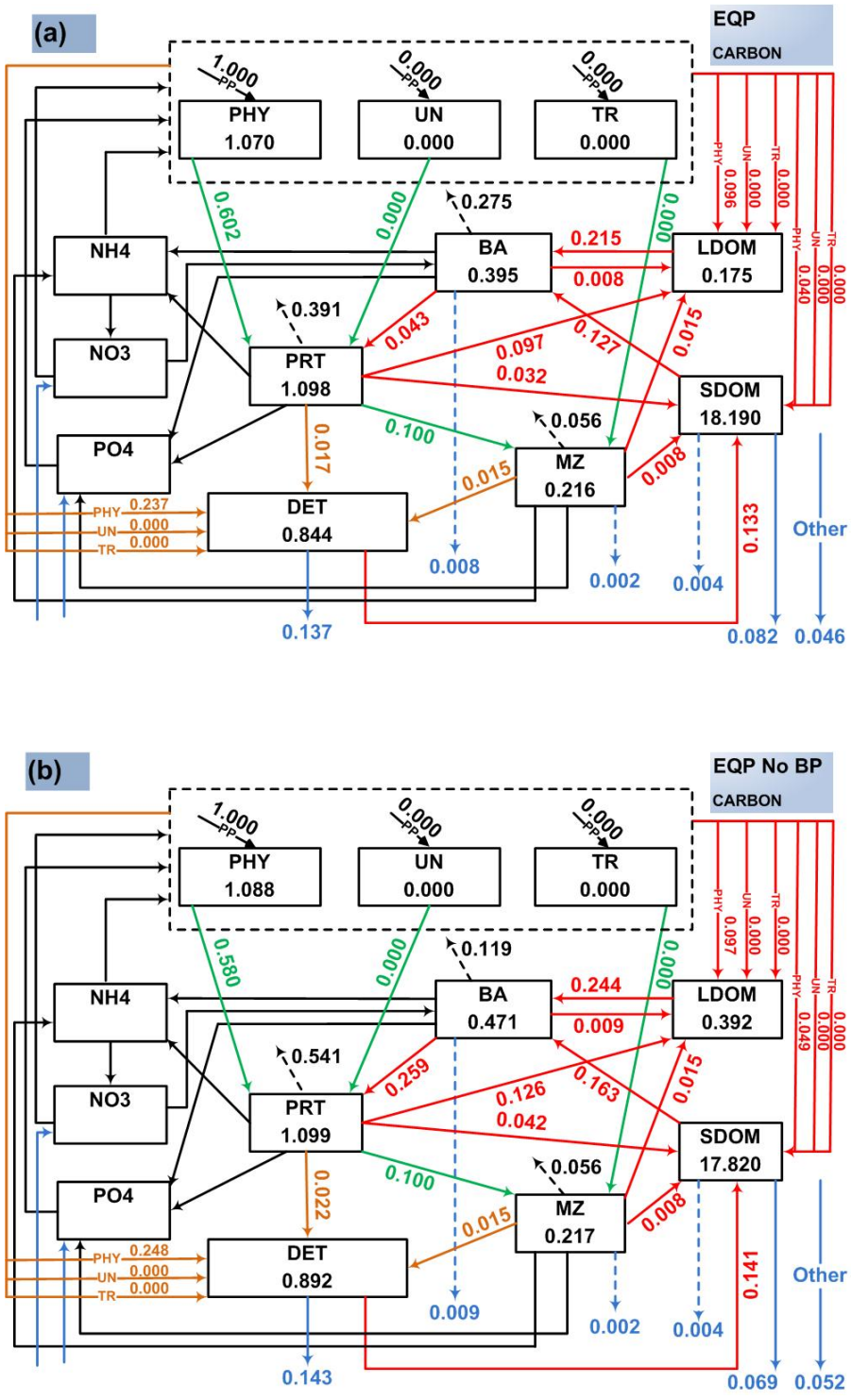

(a)
EQP
CARBON
PHY
1.070
UN
0.000
TR
0.000
1.000
0.000
0.000
NH4
NO3
PO4
BA
0.395
LDOM
0.175
PRT
1.098
SDOM
18.190
MZ
0.216
DET
0.844
0.275
0.391
0.056
0.602
0.000
0.000
0.215
0.008
0.043
0.097
0.032
0.127
0.015
0.100
0.017
0.015
0.008
0.237
0.000
0.000
0.008
0.002
0.004
0.133
0.137
0.082
0.046
Other
(b)
EQP No BP
CARBON
PHY
1.088
UN
0.000
TR
0.000
1.000
0.000
0.000
NH4
NO3
PO4
BA
0.471
LDOM
0.392
PRT
1.099
SDOM
17.820
MZ
0.217
DET
0.892
0.119
0.541
0.056
0.580
0.000
0.000
0.244
0.009
0.259
0.126
0.042
0.163
0.015
0.100
0.022
0.015
0.008
0.248
0.000
0.000
0.009
0.002
0.004
0.141
0.143
0.069
0.052
Other

Figure 2 *Continued*

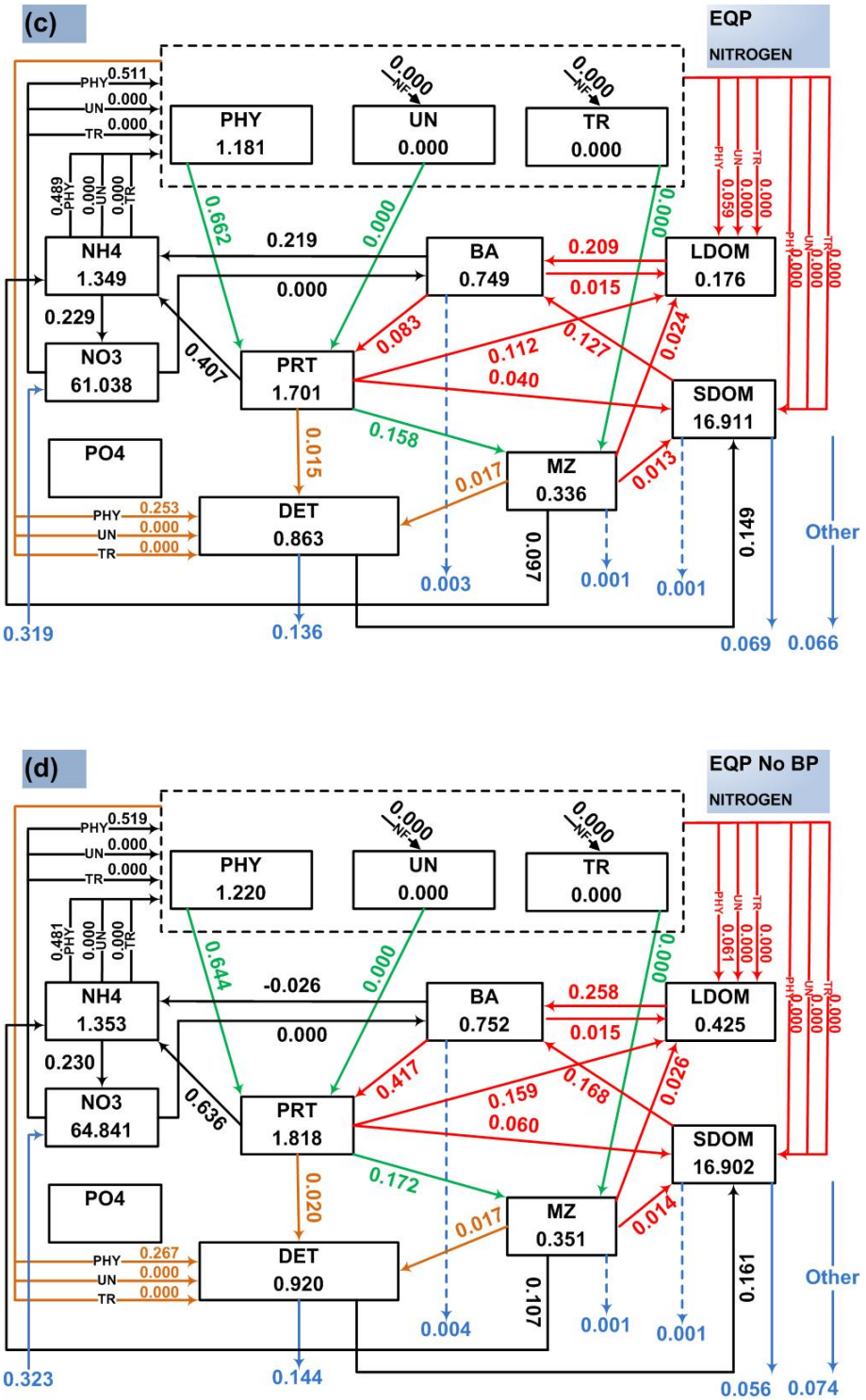


Figure 3

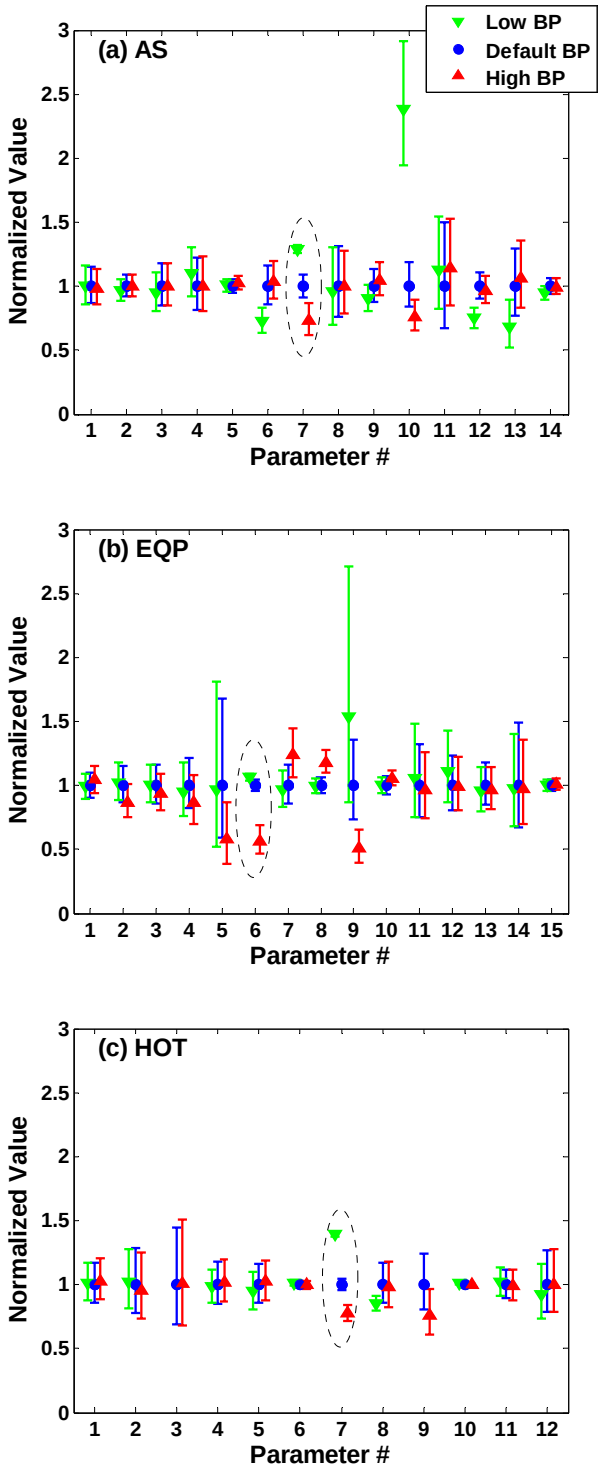
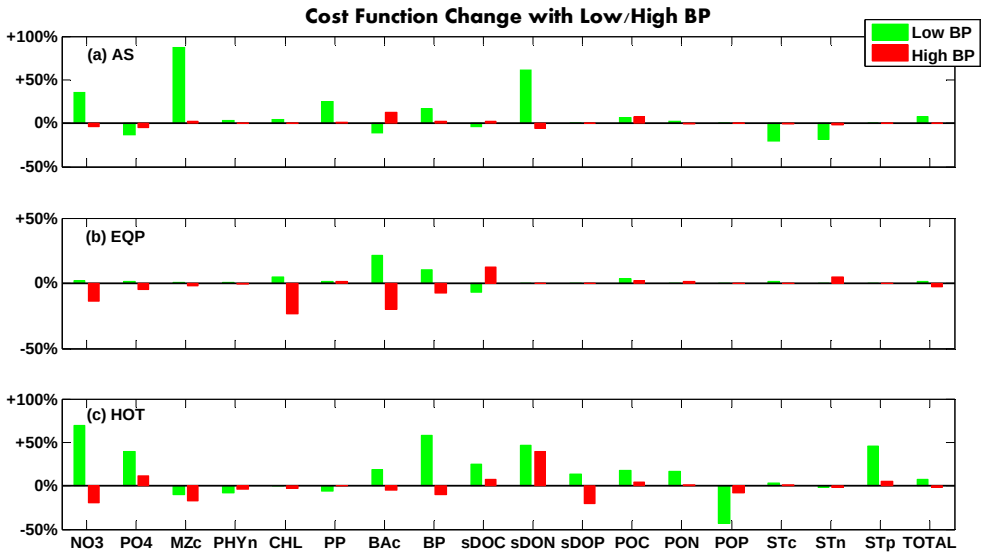

Low BP
Default BP
High BP
(a) AS
Normalized Value
Parameter #
(b) EQP
Normalized Value
Parameter #
(c) HOT
Normalized Value
Parameter #

Figure 4


Cost Function Change with Low/High BP
Low BP
High BP
(a) AS
(b) EQP
(c) HOT
+100%
+50%
0%
-50%
NO3
PO4
MZc
PHYn
CHL
PP
BAc
BP
sDOC
sDON
sDOP
POC
PON
POP
STc
STn
STp
TOTAL

Chapter 6. Summaries and Conclusions

1. Biological Model

Prokaryotic microbes are essential components of the ecosystems in the open ocean. *Prochlorococcus* and *Synechococcus* cyanobacteria dominate autotrophic production in many areas (Glover 1986, Burkill et al. 1993, Campbell 1994, Liu et al. 1997, Liu et al. 1998). Another group of cyanobacteria, *Trichodesmium* spp., is one of the major autotrophic dinitrogen (N_2) fixers (diazotrophs) and introduces a significant amount of new bio-active nitrogen into the open ocean (Karl et al. 2002, Capone et al. 2005, Westberry & Siegel 2006). A newly identified autotrophic unicellular diazotrophic cyanobacterium may be equally important as *Trichodesmium* spp. in fixing N_2 (Zehr et al. 2001, Montoya et al. 2004). Recent studies showed these diazotrophs have a critical role in supporting oceanic new production (Carpenter & Romans 1991, Gruber & Sarmiento 1997, Karl et al. 1997). The dynamics of heterotrophic bacteria can significantly impact carbon cycling and nutrient regenerations in the ocean. Heterotrophic *Bacteria* and possibly *Archaea* (termed heterotrophic bacteria hereafter) are the two major consumers of dissolved organic matter (DOM) (Kirchman 2008), the dominant organic pool in the open ocean (Hedges 1992). About 50% of primary production (PP) by phytoplankton in oceanic ecosystems is released as DOM (Ducklow & Carlson 1992, Williams 2000). Most of the organic carbon production metabolized by bacteria is respired to CO_2 (del Giorgio & Cole 1998, Robinson 2008) along with the regeneration of inorganic nutrients (Kirchman 2000, Landry 2001). Prokaryote biomass can be grazed by protists and then by higher trophic levels: a process called the “microbial loop” (Azam et al. 1983). In this

dissertation, an ecosystem model was constructed focusing on microbial components in the ocean, from autotrophic non-diazotrophs, *Trichodesmium* spp. and unicellular diazotrophs, to heterotrophic bacteria and DOM.

The model has several novel features. The flexible stoichiometry of the modeled elements (carbon, nitrogen and phosphorus) is the fundamental structure to facilitate the controls on N_2 fixation and heterotrophic bacteria-DOM dynamics. N_2 fixation rates are determined by the difference between nitrogen demand and the existing supply of bioavailable nitrogen (ammonium and nitrate), and also by the available carbon to support energy requirement. The two groups of diazotrophs have different light-dependence and grazing pressure. The DOM pool is separated into labile and semilabile DOM which have different dynamics in terms of their consumption by heterotrophic bacteria. The semilabile DOM has variable lability determined by its nutrient content. The heterotrophic bacteria can select semilabile DOM to take up more nutrients than the average bulk pool. The heterotrophic bacterial growth efficiency varies with their production.

2. Data Assimilating Model

To integrate the available observations with the model, a data assimilating framework using a variational adjoint method was constructed to optimize the model parameters and minimize the differences between the observations and model results (Lawson et al. 1995, Friedrichs et al. 2007). Unlike traditional ways to tune individual model parameters, data assimilation is more objective and efficient. It looks for the best

solutions for many parameters to fit various observations simultaneously. It can also help to identify the sensitivity of the model to each parameter and judge the sufficiency of observations in order to constrain parameters.

The data assimilating model was applied to three open ocean sites, the Arabian Sea (AS), Equatorial Pacific (EQP) and Station ALOHA at the Hawaii Ocean Time-series (HOT) covering a 1-dimensional surface 150 m deep (200 m for HOT). Those applications were supported by observations collected by the U.S. JGOFS project in 1994 for AS, 1992 for EQP and also longer-term data by the HOT program since 1988. Up to 17 types of major biogeochemical observations were assimilated including nutrients (nitrate and phosphate), phytoplankton biomass (nitrogen biomass and chlorophyll *a*), PP, heterotrophic bacterial biomass and production, mesozooplankton biomass, dissolved and particulate organic matter (carbon, nitrogen and phosphorus) and export of organic matter (carbon, nitrogen and phosphorus collected by sediment traps). Thus the model was not only being optimized, but also being strictly tested for its robustness to fit different properties of the ecosystem simultaneously.

The modeling was conducted for AS, EQP and one year of HOT for cross-comparison in different oceanic regions. A long-term model study was also conducted for HOT in 1991 – 2005 in order to study the decadal variability of the ecosystem. All the experiments showed good fit to most observations after the data assimilation, which made the modeled results more reliable and persuasive, and possibly more realistic.

3. Findings

The modeled results at all the three open ocean sites showed some consistemcies in heterotrophic bacterial activities. Although the rates of heterotrophic bacterial net production (BP) are different at these sites, the heterotrophic bacterial carbon demand (BCD), the carbon consumed by heterotrophic bacteria, equals one third to half of the PP, and the nutrients recycled by heterotrophic bacteria can support a quarter of PP. Previous studies focused on labile DOM as the major resource for BCD, while emphasizing semilabile DOM as a potential export instead of a resource for heterotrophic bacteria (Nagata 2008). The modeled results from this study show that semilabile DOM supports significant amount (11 – 37%) of heterotrophic bacterial carbon demand, and appears as a temporally-steady carbon resource compared to labile DOM. The amount of BP has been usually considered to return to the major food webs in the concept of the "microbial loop" (Azam et al. 1983). However, the model shows that BP is not completely consumed by protozoa and returned to the food web, but about one third to a half of BP goes to other pools including labile DOM resulting from viral lysis and also refractory DOM. The model also suggests a new form of export: the production of refractory DOM, mainly from bacterial biomass. This pool is considered as bio-unavailable over multiannual time scales in the surface ocean. This form of export consists of one third to a half of total export at AS and HOT. These conclusions suggested by the model are constrained directly by the observed heterotrophic bacterial biomass and production and the concentrations of DOM including its carbon, nitrogen and phosphorus components. In

addition, these features are not contradicted by the observations related to the nonmicrobial part of the ecosystem, as the model can still fit most observations simultaneously. Thus the conclusions related to heterotrophic bacteria and DOM are at least compatible to our current known understandings of oceanic ecosystems.

To address the importance of BP measurements for constraining the model and the uncertainties of conversion factors related to deriving BP rates from current radioisotope based methods, several experiments were conducted at the three sites by removing BP from assimilation, and by assimilating low and high BP estimates as supplied by the commonly-used empirical ranges of the conversion factors. The experiments showed that BP data were useful to strengthen constraints on the model parameters for correct explanations of some mechanisms of ecosystem processes. The low/high BP experiments showed that both the low BP (5% of PP) and the high BP (25% of PP) cannot fit into our current knowledge and other observations, which suggests that BP in the open ocean is probably inside the previous suggested range of 5 – 25% of PP (Ducklow 2000, Anderson & Ducklow 2001). Although certainly spatial and seasonal variations exist for the conversion factors of BP, the model suggests that currently used range probably can be applied to most habitats in the open ocean.

As biogeochemical observations accumulate at Station ALOHA as the HOT project continues, the PP at the station has showed an upward trend in the past two decades. The model created in this dissertation can successfully simulate the decadal increase over 1991 – 2005, and identified the mechanisms for two depth typical layers: 0 – 20 m and 50

– 70 m. The deepening mixed layer and increased physical nitrate supply from the deep ocean supported the increasing PP at both layers. It is the dominant mechanism for the subsurface 50 – 70 m, but two other mechanisms are equally important in the surface 0 – 20 m. First, the increased physical nitrate supply inhibits dinitrogen (N_2) fixation in the surface, and this replacement of one new nitrogen form (N_2 fixation) with another form (physical supply) enhanced the nitrogen recycling efficiency and thus the PP. It implies that although N_2 fixation and physical nitrate supply are widely considered as two major forms of new nitrogen supply (Karl et al. 2002), they are not equally important in terms of fueling primary production. Second, the consumption of semilabile DOM in the surface decouples and exceeds its biological production when the temporal change of these two terms are compared, which identifies semilabile DOM as another new form of nitrogen supply on the decadal time scale.

References

Anderson TR, Ducklow HW (2001) Microbial loop carbon cycling in ocean environments studied using a simple steady state model. Aquat. Microb. Ecol. 26:37-49

Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad F (1983) The ecological role of water-column microbes in the sea. Mar. Ecol. Prog. Ser. 10:257-263

Burkill PH, Leakey RJG, Owens NJP, Mantoura RFC (1993) Synechococcus and its importance to the microbial foodweb of the northwestern Indian Ocean. Deep Sea Research Part II: Topical Studies in Oceanography 40:773-782

Campbell L, H. A. Nolla, and D. Vaulot (1994) The importance of Prochlorococcus to community structure in the central North pacific Ocean. Limnol. Oceanogr. 39:954-961

Capone DG, Burns JA, Montoya JP, Subramaniam A, Mahaffey C, Gunderson T, Michaels AF, Carpenter EJ (2005) Nitrogen fixation by Trichodesmium spp.: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. Global Biogeo. Cycles 19:GB2024

Carpenter EJ, Romans K (1991) Major role of the cyanobacterium Trichodesmium in nutrient cycling in the North Atlantic Ocean. Science 254:1356-1358

del Giorgio PA, Cole JJ (1998) Bacterial Growth Efficiency in Natural Aquatic Systems.

Ann. Rev. Ecol. Syst. 29:503-541

Ducklow HW (2000) Bacterioplankton production and biomass in the oceans. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley, New York, p 85-120

Ducklow HW, Carlson CA (1992) Oceanic bacterial productivity. Adv. Microb. Ecol. 12:113-181

Friedrichs MAM, Dusenberry JA, Anderson LA, Armstrong RA, Chai F, Christian JR, Doney SC, Dunne J, Fujii M, Hood R, McGillicuddy DJ, Moore JK, Schartau M, Spitz YH, Wiggert JD (2007) Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. Journal of Geophysical Research-Oceans 112:C08001, Doi 08010.01029/02006jc003852

Glover HE, L. Cambpell, and B.B. Prezelin (1986) Contributions of Synechococcus to size fractionated primary production in three water masses in the Northwest Atlantic Ocean. Mar. Biol. 91:193-203

Gruber N, Sarmiento JL (1997) Global patterns of marine nitrogen fixation and denitrification. Global Biogeo. Cycles 11:235-266

Hedges JI (1992) Global biogeochemical cycles: progress and problems. Mar. Chem. 39:67-93

Karl D, Michaels A, Bergman B, Capone D, Carpenter E, Letelier R, Lipschitz F, Paerl H, Sigman D, Stal L (2002) Dinitrogen fixation in the world's oceans. Biogeochemistry 57/58:47-98

Karl DM, Letelier R, Tupas L, Dore J, Christian J, Hebel D (1997) The role of nitrogen

fixation in biogeochemical cycling in the subtropical North Pacific Ocean. Nature 388:533-538

Kirchman DL (2000) Uptake and regeneration of inorganic nutrients by marine heterotrophic bacteria. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, Inc., New York, p 261-288

Kirchman DL (2008) Introduction and overview. In: Kirchman DL (ed) Microbial Ecology of the Oceans (second edition). John Wiley & Sons, Inc., Hoboken, New Jersey

Landry M (2001) Microbial loops. In: Steele JH, Thrope SA, Turekian KK (eds) Encyclopedia of Ocean Sciences. Academic Press, London, p 1763-1770

Lawson LM, Spitz YH, Hofmann EE, Long RB (1995) A data assimilation technique applied to a predator-prey model. Bulletin of Mathematical Biology 57:593-617

Liu H, Campbell L, Landry MR, Nolla HA, Brown SL, Constantinou J (1998) Prochlorococcus and Synechococcus growth rates and contributions to production in the Arabian Sea during the 1995 Southwest and Northeast Monsoons. Deep-Sea Res. II 45:2327-2352

Liu HB, Nolla HA, Campbell L (1997) Prochlorococcus growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean. Aquat. Microb. Ecol. 12:39-47

Montoya JP, Houll CM, Zehr JP, Hansen A, Villareal TA, Capone DG (2004) High rates of N_2 fixation by unicellular diazotrophs in the oligotrophic Pacific Ocean. Nature

430:1027-1032

Nagata T (2008) Organic Matter-Bacteria Interactions in Seawater. In: Kirchman DL (ed) Microbial Ecology of the Oceans, Second Edition. John Wiley & Sons, Inc., Hoboken, New Jersy, p 207 - 241

Robinson C (2008) Heterotrophic bacterial respiration. In: Kirchman DL (ed) Microbial Ecology of the Oceans (second edition). John Wiley & Sons, Inc., Hoboken, New Jersey

Westberry TK, Siegel DA (2006) Spatial and temporal distribution of Trichodesmium blooms in the world's oceans. Global Biogeochem Cycles 20:GB5016

Williams PJl (2000) Heterotrophic bacteria and the dynamics of dissolved organic material. In: Kirchman DL (ed) Microbial Ecology of the Oceans. Wiley-Liss, New York, p 153- 200

Zehr JP, Waterbury JB, Turner PJ, Montoya JP, Omoregie E, Steward GF, Hansen A, Karl DM (2001) Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. Nature 412:635-638