1. DETAILED RESPONSES TO REVIEWERS

Response to Comments by Reviewer 1 (FMM Monteiro)

(Reviews are included in black font; Responses are in blue font)

The study by Kuhn et al. investigates the role of heterotrophic nitrogen fixers in the Gulf of Aqaba. To do so, they develop the first ocean model of heterotrophic nitrogen fixers in a 1D setting which also includes other main types of nitrogen fixers. They optimise their model parameters with time series observations of this region and then validate their model's results looking at different model's version (with/without heterotrophic N2 fixers, with/without explicit N2 fixation . . .) to look at the sensitivity of the different types of nitrogen fixers on the ocean biogeochemistry. They find that heterotrophic N2 fixers are key in representing observed concentration of nitrate and oxygen in the deep ocean.

This is an important study as the first model study to include heterotrophic nitrogen fixers, a model which has been carefully optimised and validated with an extensive time series. The paper is very well-written, clear, concise, and presents well-designed modelling experiments. I thus strongly encourage its publication. My only concern is on the conclusion in relation to N* as the model doesn't capture very well the observations (see main comments below).

Response: We are grateful for the positive assessment and appreciate the constructive comments. As described in our detailed responses below, we have modified the manuscript accordingly and believe it has improved significantly as a result.

Main comments

N* model reproduction While the model (H3 in particular) does a very good job at representing NO3 and O2 observed concentrations, the model seems to be quite off from the N* data. This is not well enough highlighted in the paper which currently presents heterotrophic nitrogen fixers being key on reproducing the N* values. Because the model (even H3) is not able to capture most of the observed variability in N*, I don't think the model results support well enough this conclusion. I wonder what could be missing in the model and if you have thought about the role of preferential P remineralisation on N*. I wrote a paper in 2012 investigating the role of preferential remineralisation of P on the distribution of N* in the North Atlantic. In that region, this mechanism

is necessary to reproduce the observed sub-surface maximum of N*. Here your model seems to have too small values in PO4 and NO3 at depth. Would it possible to test in your model the effect of preferential remineralisation of P to see if that helps re- producing the observed N* variability? If not, at least mention it. In our 2012 model, preferential remineralisation of P helped to get higher concentration of P, enhanced nitrogen fixation and then resulted in higher N* value at depth (as N* increases after the remineralisation of diazotrophic matter).

Monteiro, F. M., & Follows, M. J. (2012). On nitrogen fixation and preferential reminer- alization of phosphorus. Geophysical Research Letters, 39(6).

Response: Thank you for suggesting this paper. We included this citation in section 5.4, where we discuss uncertainties and limitations of our study:

"*…preferential PO4 remineralization may more directly increase deep-PO4 concentrations as organic matter is decomposed in the bottom layers. The inclusion of this process in a model of the North Atlantic Ocean improved the representations of biogeochemical characteristics of the area and increased the N_2 fixation rates obtained by the model (Monteiro and Follows, 2012).*"

Nevertheless, as suggested by other reviewers, we moved our focus away from N* by removing figures 3 and 6 from the manuscript.

Minor comments

P2, Line 1: Need to mention about atmospheric N sources

Response: We agree. The sentence was changed to:

"Locally the supply of new nitrogen can occur through several mechanisms, including microbially mediated N_2 fixation, diapycnal mixing injecting deep nitrate (NO3) into the surface, lateral transport, ***atmospheric N sources*** *and riverine input."*

P3, Line 5: Need to justify in the introduction why the Gulf of Aqaba is an interesting region to study nitrogen fixation.

Response: Agree. The referred paragraph in the introduction was changed to:

"*In this study we explore the biogeochemical signatures that result from different assumptions about the ecological niches occupied by diazotrophs.* ***Our study area is the Gulf of Aqaba, a northern extension of the Red Sea. Aside from the reported presence of diverse diazotrophs types,***

the morphology of the Gulf of Aqaba limits horizontal transport of deep waters, thus allowing us to simplify the physical model's complexity of the model and focus on the biological component."

P3, Lines 21-23: Sentence not clear. Can you amend?

Response: We simplified this sentence for clarity". The modified sentence reads:

"*Since inflow is restricted to warm surface waters, the Gulf's deep water masses (>300 m) are locally formed (Wolf-Vecht et al., 1992; Biton et al., 2008) and have negligible horizontal transport toward the exterior (Klinker et al., 1976; Manasrah et al., 2006).*"

P5, Line 16: The model description is confusing on if/how H0 and H0' represent N2 fixation. Here you say "without explicit N2 fixation . . . and follows the model equations described in Fennel et al. (2006, 2013)", which would mean that there is a non explicit representation of N2 fixation in the model. If so, replace with "with non-explicit N2 fixation representation" and then describe briefly how it is done in Fennel's papers. But later on there are many references to "with no N2 fixation" (P5/Line 18, Figure 4, P10 Line 5, P14 Line 20, . . .). Can you amend accordingly?

Response: We modified the description as follows:

"*H0 is the base model without diazotrophic plankton groups and follows the model equations described in Fennel et al., (2006, 2013). In general, the model follows Monod kinetics using a fixed N:P ratio ($R^{nf}_{N:P}$= 16). Sensitivity to the constant N:P ratios is explored in section 4.2.3. We test the H0 model with and without a sediment denitrification flux (model versions H0 and H0', respectively). H0 includes denitrification but no N_2 fixation, as no diazotrophs are considered. H0' does include neither denitrification, nor diazotrophic organisms, and thus the underlying assumption of this model version is that there is a balance between inputs from N_2 fixation and losses of fixed nitrogen due to denitrification. When present, the denitrification flux follows Fennel et al. (2013) with a loss fraction 6 mol N_2 per mol of organic matter remineralized at the sediment-water interface. This generates an average sediment denitrification flux of 0.25 ± 0.46 mmol N m^{-2} d^{-1}, with a maximum value of 3.01 mmol N m^{-2} d^{-1}.*"

P6, Line 5: Isn't there any evidence of DDA in this region or atmospheric source of N? Please add comments as they can be potential important sources of N.

Response: We extended the discussion about other potential sources of nitrogen (section 5.4):

"*There are other sources of nitrogen that were not explored in the present study and we discuss*

here briefly. For instance, we did not include contributions to N_2 fixation by diatom-diazotroph associations, which are significant in other regions. While the genetic material of diatom-diazotroph associations has been detected in the Gulf of Aqaba, it is not as abundant as unicellular diazotrophs, Trichodesmium and proteobacteria (Kimor et al. 1992, Foster et al., 2009). In general, due to the oligotrophic characteristics of the region, small phytoplankton species (<8 μm) contribute more than 90% of the chlorophyll-a standing stock (Lindell & Post 1995, Yahel et al. 1998). Dinoflagellates and diatoms together correspond to less than 5% of the phytoplankton biomass, except during ephemeral diatom blooms during spring when they can account for nearly 50% of the total biomass (Al-Najjar et al., 2006).

Another source of nitrogen that has received interest in this region is atmospheric deposition, as the Gulf receives considerable dust input from the surrounding deserts. Recently, it has been shown that atmospheric dust input does not correlate with chlorophyll variability in surface waters of the Gulf of Aqaba (Torfstein and Kienast, 2018). A previous study suggested that atmospheric deposition of nitrogen could support over 10% of surface primary production in the region, based on measurements of local aerosol composition and a dust deposition model (Chen et al., 2007). However, this estimate had a relatively large uncertainty due to errors associated with the deposition flux calculation and the temporal variability in dust flux (Chen et al., 2007). Moreover, very low nitrogen concentrations and N:P ratios lower than Redfield from the surface down to 80 m were observed during the same time period (Foster et al., 2009). Therefore, the role of atmospheric nitrogen inputs remains uncertain."

P7, Line 2: I could not find how denitrification is represented in the model. Can you make sure it is described?

Response: The denitrification representation was described on the supplement. Now we have included it in the methods sections:

"*When present, the denitrification flux follows Fennel et al. (2013) with a loss fraction of 6 mol N_2 per mol of organic matter remineralized at the sediment-water inetrface. This generates an average sediment denitrification flux of 0.25 ± 0.46 mmol N $m^{-2} d^{-1}$, with a maximum value of 3.01 mmol N $m^{-2} d^{-1}$.*"

P9, Line 10: Need to comment on the model performance for PO4. Especially for 100-600 m where H3 matches well observations for NO3, still cannot get PO4 right.

Response: We now state more clearly that: "*All model versions represent similar vertical distributions of PO_4 and underestimate its deep-water concentrations by the end of the series*" and we discuss this in the uncertainties and limitation section.

P9, Line 11-12: Can you refer to Figure 3 here?

Response: The previous Figure 3 was removed. Figure 3 now refers to a different result.

P9, Lines 20-24: This paragraph does not add much so I would be inclined to remove it.

Response: This paragraph was deleted.

P10, Line 9: It would be good to add a statement on why heterotrophic N2 fixers improve the representation of deep oxygen as a key result of your paper.

Response: We now put more emphasis on oxygen in the discussion and added the following text:

"*This model improves the representation of NO_3 and O_2 at depth (Figures 3, 7). Changes in deep NO_3 can be explained through the enrichment of detritus, while changes in O_2 occur because the heterotrophic group becomes an additional sink of O_2 at depth*."

P10, Line 18-20: Need to comment on potential reasons why H3 accumulates NO3 over time.

Response: We now state this more clearly:

"*All models that consider nitrogen fixation accumulate nitrogen at different rates, as they enrich the nitrogen content of detritus, which is then remineralized at depth over time*."

P11, Line 3-8: Need to add comments on why H3 and H2 have much higher N2 fixation rate than observations between 0-DCM in the Summer 2010.

Response: In the H3 and H2 models the large N_2 fixation rates during summer are due to the contribution of blooming Trichodesmium. This is likely to be overestimated under the current model configuration and requires further calibration as more information becomes available to verify the seasonal cycle of nitrogen fixation in the region. We are acknowledging this by including the following sentence:

"*… massive blooms are rare in the Gulf of Aqaba (Foster et al., 2009; Mackey et al., 2007) and the model probably overestimates the contribution of Trichodesmium spp.'s annual blooming to*

total N_2 fixation rates, as seen in the much larger surface N_2 fixation rates generated by H2 and H3."

P11, Line 17-18: I would be more subtle about the models' abilities to replicate N* as it is still quite far from the observations (see my main comments above).

Response: Following this and other reviewer's comments we removed the figures and discussion concerning N*.

P12, Lines 17-32: While this is an interesting section about the contribution of N2 fixation on PP, why include results from H3a which is not as realistic as H3?

Response: Following the Reviewer's recommendation, this text was removed.

P13, Section 5.3: One of the main points of the paper is to highlight the important role of heterotrophic nitrogen fixers. I feel this section could then be a lot stronger highlighting all the effect of heterotrophic N2 fixers on the ocean biogeochemistry of this region. Here for instance I would mention that heterotrophic N2 fixers improve the NO3 and O2 concentrations at depth, as well as the contribution of heterotrophic bacteria to total N2 fixation. Also, would it be possible to plot the model difference between H2 and H3 to show the impact of heterotrophic N2 fixation?

Response: We have modified the text accordingly. In term of plots, we have decided to show only the difference in terms of total N_2 fixation, which presented in Figure 10b.

Figure 10 is not mentioned in the main text. It looks interesting so probably worse describing it at some point.

Response: Thank you for pointing this out. Figure 10 is now referred to in sections 5.2 and 5.3

Response to Comments by Reviewer 2

(Reviews are included in black font; Responses are in blue font)

GENERIC COMMENT:

This paper presents a model for heterotrophic nitrogen (N) fixation, implements it in a 1D context in the Arabian Sea and uses the model to test hypotheses on the relative contributions to N fixation by the different organisms. The subject fits perfectly in the Journal remits, and the work is highly relevant and it will be an important contribution to the topic of N fixation. I particularly liked the

use of genetic algorithm for calibration and use of the model to test hypotheses. Authors set up a generally good framework to perform those test, unfortunately I believe that some further tests are needed in order to properly attribute the changes in the model outputs to the N fixation trait (see main comments).

Response: We are grateful for the positive assessment and appreciate the constructive comments. As described in our detailed responses below, we have modified the manuscript accordingly and believe it has improved significantly as a result.

MAIN COMMENTS: First of all, I would strongly encourage authors to be more comprehensive in the model description the supplement, because some key details are not clear, particularly on how the equations change in the different model set-up. In particular, in model H2 how the equation for zooplankton growth changes? Does Zooplankton see a single pool of phytoplankton formed by non-fixers and unicellular fixers, or does it graze separately on both? Given the non-linearity of the limitation function, the two options are very different. I would suggest writing explicitly all equations of H1, H2 and H3 that differ from H0 instead of summarising with sentences like "All other state variable equations are modified accordingly"

Response: In response to this suggestion we have now included the complete sets of equations in the supplements and have described them in more detail to clarify the model assumptions.

The main concern is that authors directly compare models with very different structure and then attribute all changes observed in the results to the process without separating the impact of the biogeochemical process from the impact of the different model structure. For instance, in H3 authors added the heterotrophic nitrogen fixers, by adding to the implicit first-order mineralisation scheme of the small detritus, a more dynamic one that includes an explicit heterotrophic group (Hf). Such a big change in the model structure is bound to profoundly impact the model results, regardless of the N fixation ability of the heterotrophic group, because the whole dynamic of mineralisation is changed. I would recommend the authors to implement a H3' model where a non-fixer group of heterotrophic organisms uses organic and inorganic for of both N and P is used as Hf.

The comparison between H0 and H3' would enable to understand how much of the mismatch between simulated and observed bottom waters N and O2 is due to an underestimation of mineralisation, while comparing H3' with H3 will allow to assess how the N fixation trait influence those dynamics. The comparison of H0 and H3' is much more important because the mineralisation

rates have not been calibrated, and therefore could be affected by an initial bias.

Response: We find this to be an excellent suggestion and have performed the additional simulation as suggested. In the model description we have added:

"*An intermediate version H3' is used as a control, where the heterotrophic organisms do not fix nitrogen and are limited by the availability of nitrogen in inorganic forms and from small detritus. Model H3 eliminates the nitrogen limitation and the heterotrophic group becomes a heterotrophic diazotroph group.*"

The new results are included only in Figures 3 and 7, and described in section 4.2 "Model Results" to preserve the clarity of other plots. This additional simulation did not affect our main conclusions.

Similarly, when comparing model with 1 phytoplankton group (H0) with models with multiple PFT, all trophic dynamics can change, due to non-linearity in the graz- ing. Since Zooplankton dynamics are not shown, nor detailed equation for grazing and zooplankton growth in the different models, it is impossible to me to assess if the implementation of H1' and H2' similar to H3' are to be recommended or not.

Response: While we have tried to keep the model structures as consistent and comparable as possible, we acknowledge that comparing models with different complexities is challenging due to the various changes that may occur as consequence of changing trophic structure. We have expanded the model equations in the supplement to improve the clarity. However, further examination of the effects of different grazing functional forms, or other similar exercises are beyond the intended scope of the study.

Another main comment is related to the Redfieldian assumption. While I fully acknowledge the long tradition of Redfield ratio based models and data analyses, their power and their advantage, I'm always a bit concerned when these are used to draw conclusions on nutrient ratio dynamics, particularly in the short temporal and spatial scales. Phytoplankton internal nutrient ration and nutrient uptake are far from being constant and fixed to the Redfield ratio and they also varies a lot from species to species (e.g. Geider and La Roche, European Journal of Phycology, 2011, http://dx.doi.org/10.1017/S0967026201003456). I appreciate that this complexity is impossible to fully reproduce in biogeochemical model and therefore the Redfield assumption can still be used as first order approximation in simple biogeochemical model, however I would not use those model to analyse the instantaneous dynamic of nutrient ratios because this will be strongly affected by the

huge assumption of fixed stoichiometry. Figure 6 itself shows how the model is not able to capture the wide variability of DIN:DIP ratio. For this reason, I would suggest to cut the part related to N*, or alternatively, repeat the analysis using annual means of DIN and DIP and include a discussion on the importance of non-Redfieldian dynamics.

Response: Following this and other Reviewers' comments we have decided to remove figures and discussion concerning N*. Also, sensitivity experiments to N:P ratios now illustrate the effect of changes in non-fixing phytoplankton and diazotroph stoichiometry (see responses to Reviewer 3).

SPECIFIC COMMENTS: Page 5, lines 19-20: in H0' N fixation and denitrification are balanced: where did denitrification occur? In the benthos? I recommend adding some detail to better interpret the vertical dynamics simulated by this model implementation

Response: The denitrification representation was included in the model description in the supplement: *"When present, the denitrification flux follows Fennel et al. (2013) with a loss fraction of 6 mol N_2 per mol of organic matter remineralized"*. We added this descriptive sentence in the methods section of the main text.

Section 4.2.2.: Top left panel of figure 5 shows that model H2 and H3 are significantly overestimating surface nitrogen in the last 4 years of calibration, with the exception of the deep mixing events in winter 2007/2008. H3 largely overestimates surface nitrogen also in the validation period (figure 8). This important dynamic is not discussed in the paper.

Response: Geomorphology and bathymetry limit water exchange fluxes between the Gulf of Aqaba and the Red Sea to the upper 300 m. It this unlikely that horizontal transport could explain the observed accumulation in deep NO_3, but exchange of surface waters during summer months does occur. A possible explanation is that some of the DIN produced should be exported to the outside of the Gulf, which our 1D model does not account for. The following sentence was added to section 5.4:

"*It is, therefore, unlikely that horizontal transport could explain the observed accumulation of deep NO_3. Nevertheless, transport of nitrogen-enriched sub-surface waters from the Gulf of Aqaba towards the exterior may dampen the long-term accumulation of nitrogen*."

Section 4.2.4.: while I agree that H3 better compares with observed deep values, in the last couple of years a significant trend in deep nitrogen appears in the simulation and it's not in the data. I

suggest authors to comment on that.

Response: Please see previous response concerning export of DIN to outside the Gulf.

TECHNICAL COMMENT: Page 12, line 18: in 10b, the dot corresponding to Capone and Carpenter 1982 shows a N fixation equal or close to 0, that is quite different from the values simulated by the different flavour of H3 Figures 2,4,7: I recommend the authors to redraw the picture using a perceptually uniform and colour-blind friendly colourmap like viridis, inferno, magma or plasma in Python or Parula in Matlab. More details on the importance of this in the following video https://www.youtube.com/watch?v=xAoljeRJ3lU

Response: Taken, we have redrawn the data figure in the color-blind friendly colormap "parula".

Response to Comments by Reviewer 3

(Reviews are included in black font; Responses are in blue font)

The authors present a study which compares a series of 1-D biogeochemical models of increasing complexity with respect to the representation of different diazotrophic organisms against an in situ time series data set from a location in the Gulf of Aqaba in the Red Sea. Comparison of the output of these models with the in situ data and specifically the nutrient concentrations and ratios/differences between N and P con- centrations (as quantified by the derived N* variable) is subsequently used to argue for a substantive contribution by heterotrophic diazotrophs to N2 fixation within the region. The manuscript is well written and in general the study rationale and model experiments appeared well designed (although see specific comments below). The subsequent results and potential implications of the study were certainly interesting and overall I felt there was much of value within the study. However, as outlined below, I have a few concerns I would like to see the reviewers address.

Response: We are grateful for the overall positive assessment and appreciate the constructive comments, which we respond to in more detail below. We believe the resulting edits have greatly improved the manuscript.

Major comment:

The authors appear to undertake a thorough job of optimising many of the parameters related to their model (see e.g. Page 6 and Table 2). However, given the key question(s) being addressed

within the study I was somewhat surprised that potential variability in what I would consider to be the key parameters in dictating how the diazotrophs interact with the N and P cycles were fixed, with no exploration of potential variability in these parameters. Specifically, the values of the N:P ratio within both the non-diazotrophs and the diazotrophs were fixed at 16:1 (Page 1 of Supplm.) and 45:1 (Page 5 of Supplm.) respectively. In contrast it is now fairly well recognised not only that N:P ratios within organic material can vary (see e.g. Martiny et al. 2013 Nature Geo. 6 279-283) but also (and crucially within the current context), that inferences of N2 fixation rates and interactions between diazotrophy and the cycling of N and P are highly dependent on both assumed values of these ratios and any variability within these (Mills and Arrigo 2010 Nature Geo. 4 412-416; Weber and Deutsch 2012 Nature 489 419-422). Consequently, I would suggest the authors should at least consider the implications of their assumed fixed N:P ratios for their interpretation and conclusions and perhaps also consider performing some sensitivity analysis around these currently fixed assumptions.

Response: This is a very valuable point. We performed a sensitivity analysis in which we have increased and decreased the fixed N:P ratios in the non-fixing and N fixing organisms. Section 4.2.3. and figures 5 and 6 address these experiments. Overall, changes in N:P ratios affect PO_4 concentrations more strongly than NO_3 concentrations.

Specific comments:

Page 3, Line 2: It would be worthwhile directly stating the laboratory based studies considered here were specific to Trichodesmium. As far as I am aware we have little information on how other groups might be expected to respond to the drivers men- tioned.

Response: Agree, the laboratory experiments cited consider only *Trichodesmium*, while the model study of Dutkiewicz et al. (2015) considers a set of generic diazotrophic organisms various sizes and growth rates. The following line was added:

"To our knowledge, these laboratory experiments have only explored the reaction of Trichodesmium and less information is available about the effects of climate trends on other diazotrophic organisms."

Page 5, Line 21: '. . . a generic autotrophic diazotroph. . .'

Response: Thank you for pointing out this typo.

Page 6, Lines 10-25: I was unclear whether this parameter optimisation method was performed for each of the models (H0 – H3 etc) independently or a single parameter set was used? Additionally, see major comment above, did the authors consider using the parameter optimization method for the non-diazoptroph and diazotroph N:P ratios? See also Page 14, Lines 11-16, I was unsure why this choice was made, it appears to be a big assumption within the current context.

Response: The optimization was carried out independently for each of the models. We clarified by modifying the introductory statement:

"Parameter optimization refers to the minimization of misfit between model and observations by adjusting model parameters. We applied the method first to systematically calibrate the most sensitive parameters of H0 and then to independently re-calibrate parameters in H1 to H3 after the introduction of diazotrophs."

Page 7, Line 13: maximum reported growth rates for Trichodesmium are actually >0.5 d-1, see Hong et al. (2017) Science 356 527-531

Response: Thank you for pointing us to this recent reference. We added it and modified the text accordingly in the revised manuscript.

Page 8, Line 9: see major comment above. Either this assumption should be justified, or, preferably I would suggest, some effort could be made to perform a sensitivity analysis of how the assumption influences the results/conclusions.

Response: Agree, see response above.

Page 8, Line 14: again related to comments above, some speculation on how this happens would seem appropriate in the context of this study. As a suggestion, uptake at high N:P ratios by non-diazotrophs might be one potential mechanism for shifting from inputs of nutrients with an apparent 'excess' N (i.e. positive N*) to an apparent deficit (negative N*), see e.g. Mills and Arrigo (2010).

Response: Following reviewers suggestions, we have removed N* figures as well as the text describing it. This line is no longer in the text.

Page 9, Line 19 (also Page 12, Lines 7-8): it is notable that even the most complex model struggles to reproduce the observed range in N* and I wondered whether the restriction placed on the models through the assumption of the fixed N:P ratios may be responsible for this?

Response: That is indeed a possibility that we had not explored. As brought up by Reviewer 1, other alternatives include horizontal physical transport in and out of the domain, phytoplankton stoichiometry, atmospheric deposition, and preferential PO_4 remineralization. The latter may particularly affect PO_4 at depth, while transport may affect surface values during summer (when exchange of surface waters with exterior waters has been reported to occur). We have extended our discussion about limitations and uncertainties to acknowledge these considerations. Nevertheless, given the reservations about our discussion of N* that were expressed by several of the Reviewers, we have decided to remove this figure and other references to N* in the text.

A final general point which is also related to many of those above, within the context of this study I felt that some of the important details relating to the model which were presented within the supplement would be more appropriately outlined within the main body of the text as they are likely fundamental to interpretation.

Response: The following text was added to the description of the base model to improve clarity:

"H0 is the base model without diazotrophic plankton groups and follows the model equations described in Fennel et al., (2006, 2013). In general, the model follows Monod kinetics using a fixed N:P ratio ($R^{nf}_{N:P}$=16) for phytoplankton and zooplankton. We test the H0 model with and without a sediment denitrification flux (model versions H0 and H0', respectively). H0 includes denitrification but no N_2 fixation, as no diazotrophs are considered. H0' does include neither denitrification, nor diazotrophic organisms, and thus the underlying assumption of this model version is that there is a balance between inputs from N_2 fixation and losses of fixed nitrogen due to denitrification. When present, the denitrification flux follows Fennel et al. (2013) with a loss fraction 6 mol N_2 per mol of organic matter remineralized at the sediment-water interface. This results in an average sediment denitrification flux of 0.25 ± 0.46 mmol N $m^{-2}d^{-1}$, with a maximum value of 3.01 mmol N $m^{-2}d^{-1}$." As requested by Reviewer 2, we also expanded the supplement to include all model equations.

Response to Comments by Reviewer 4

(Reviews are included in black font; Responses are in blue font)

General evaluation

This ms reports a 1D biogeochemical model analysis of time-series data from the Gulf of Aqaba from 2006–2014. The authors compare the behaviour of models with different diazotroph community structures representing various combinations of autotrophic and heterotrophic

diazotrophs. While all model versions perform similarly with respect to surface chlorophyll, only models with diazotrophy can reproduce observed nutrient (N:P) ratios and heterotrophic diazotrophy is required to explain the vertical structure of nutrient and O2 concentrations.

In general, I find this study somewhat unconvincing. The model is overly simplistic in its mechanistic foundation and ignores processes I consider essential for this kind of analysis. While I do not dispute the potential importance of heterotrophic diazotrophy for marine biogeochemistry, the conclusions and particularly the title appear overly optimistic and not well justified. The ms also appears to have been prepared rather sloppily and not thought through. The main problem is that all diazotroph parameters are unconstrained by the data, which, as outlined below, may be a consequence of the overly simplistic nature of the model or of an inappropriate cost function. Thus, in order to turn this ms into a useful contribution, the model or the cost function (or both) must be redesigned so as to achieve sensitivity to the diazotroph-related parameters.

Response: We respond one-by-one to each of the critical points that the Reviewer raises:

With regard to the comment that "the model is overly simplistic," we would like to state that our analysis is exploratory and focused on the influence of N_2 fixation, hence complexity in other parts of the ecosystem, although necessary in any global application, is intentionally minimized here. The Gulf of Aqaba is an oligotrophic system and our relatively simple model structure is able to capture the observed variability.

With regard to the title being "overly optimistic and not well justified," we have changed the title according to this Reviewer's suggestion (see response to detailed comment below).

With regard to the manuscript being "prepared rather sloppily and not thought through," we would like to point out that some of the information the Reviewer is requesting (i.e. literature sources of parameter values) is actually provided in the manuscript (Tables 2 and 3) and discussed in the text for diazotrophic organisms (section 3.3.2). In our revision we incorporated the additional information that the Reviewer requested, i.e. more details on the sensitivity experiments in the supplement, and initial conditions, more details on the ranges and literature sources for diazotrophic organisms in the main text (see detailed comments below).

With regard to "diazotroph parameters are unconstrained by the data," we would like to point out that this is not an issue with the model or the cost function. It is generally accepted among modellers that measured bulk properties like chlorophyll, nutrients and oxygen do not constrain most rates

including rates of grazing, phytoplankton mortality and, in our case, N_2 fixation (see e.g. Ward et al., 2010). No redesign of the cost function or the model will change the fact that the measured properties in the Gulf of Aqaba do not directly constrain N_2 fixation rates, simply because that information is not contained in the observations. The observations capture the impact that N_2 fixation may have only indirectly through its influence on deep-water nutrient ratios.

Ward, B.A., Friedrichs, M.A.M., Anderson, T.R., Oschlies, A., 2010. Parameter optimization techniques and the problem of underdetermination in marine biogeochemical models. J. Mar. Syst. 81, 34–43.

Specific points

1. Starting with the title, I find the wording inappropriate. While it might be possible to obtain biogeochemical evidence from a model analysis, this is certainly not the case here. I would suggest something like "Modelling heterotrophic N2 fixation ..."

Response: We have changed the title as follows:

"Modelling the biogeochemical effects of heterotrophic and autotrophic N_2 fixation in the Gulf of Aqaba"

2. Model structure. Although the authors stress that they intended to analyse mechanistic assumptions (l. 15, p. 14), I find that the model is mechanistically rather weakly founded. While simplicity is of course an important goal in model development, one must take care not to over-simplify and neglect essential processes. I think this should be at least discussed thoroughly to put the results into the right perspective. The two assumptions I find most troubling are those of (1) constant (Redfield) stoichiometry of the autotrophs and (2) obligate diazotrophy, both of which are mechanistically wrong. Fernandez-Castro et al., J. Plank. Res. 38:946 (2016), FC in the following, applied a model with variable stoichiometry and facultative diazotrophy in the subtropical North Atlantic, where the vertical distribution of N, P, and N* poses similar difficulties as in the present ms. The model of FC is otherwise very similar in structure to the present one (phytoplankton, diazotrophs, zooplankton, detritus, nutrients, DOM), so I think the differences should be discussed, particularly with respect to the relations among stoi- chiometry, export and remineralisation.

Response: We have included the citation to FC in the revised manuscript in the Introduction and

the Discussion. With regard to the particular assumptions mentioned by the Reviewer, namely constant (Redfield) stoichiometry of the autotrophs and obligate diazotrophy, we would like to comment that the overwhelming majority of models (including those used in IPCC projections) assume constant stoichiometry and, where diazotrophy is explicitly included, obligate diazotrophy. In our revision, following another Reviewer suggestion, we included an additional sensitivity experiment to show the implications of changing the stoichiometry of the diazotrophs and non-fixers. The question of obligate diazotrophy may be boil down to semantics because autotrophic diazotrophs that aren't fixing N_2 would behave like a non-diazotrophic phytoplankton, which are included in our model. We would also like to comment that our focus is not on the physiology of diazotrophs, which is analyzed in more detail by Fernandez-Castro et al. and citations therein. We acknowledge that when modeling diazotrophic cells, N allocation mechanisms are important to understand how diazotrophs are able to fixate N_2 under conditions that are traditionally thought to limit the process. However, our focus is to approach the assumptions about diazotrophs niches with a trait-based perspective. We have also changed the use of "mechanistic model" to "trait-based models" for clarity.

Comparing the parameter settings between FC and the present model, I notice a very strong discrepancy (more than a factor of 10) in the initial-slope parameter (alpha) for photosynthesis in diazotrophs, although the units are the same in both models. It is not clear from the ms how or why the very low alpha was chosen (no reference given and not optimised). But it appears to be an important parameter given that the analysis is about the vertical structure and alpha basically defines how deep in the water column autotrophic N2 fixation can occur.

Response: Our values of alpha are within the range of values typically used in ecosystem models with similar formulations (see Doney et al. 1996; Fennel et al., 2001; Fennel et al., 2002; Schartau and Oschlies 2003; Fennel et al., 2006; Moore et al., 2004). Please also notice that while FC's and our alpha parameters share the same notation, they may not refer to the same parameter because our light limitation formulations are not the same. It is not straightforward to compare these values directly to each other. Also, we would also like to note that FC's alpha values were subjectively adjusted from the original model configuration by Pahlow et al., 2013, but not optimized. FC report these parameters were adjusted to reduce the depth of N_2 fixation in their model results and even then, they obtain significant differences between the simulated and observed vertical structure of N_2 fixation. To us, these differences in parameter values simply exemplify the fact that transferring parameters measured in laboratory cultures to mathematical models representing the real ocean or in-between different models is challenging and corresponds to a different discussion in itself. Table

3 now emphasizes these references, which were only included in the text previously.

Doney, S., D. M. Glover, R. G., Najjar, 1996. A new coupled, one-dimensional biological-physical model for the upper ocean: Applications to the JGOFS Bermuda Atlantic Time- series Study (BATS) site. Deep-Sea Research II, Vol 43, No. 2-3 pp. 591-624.

Fennel, K., M. Losch, J. Schröter and M. Wenzel, 2001. Testing a marine ecosystem model: Sensitivity analysis and parameter optimization. *Journal of Marine Systems* 28/1-2, p.45-63

Fennel, K., Spitz, Y.H., Letelier, R., Abbott, M.R., 2002. A deterministic model for N_2 fixation at stn. ALOHA in the subtropical North Pacific Ocean. Deep-Sea Res. II 49, 149–174.

Fennel, K., Wilkin, J., Levin, J., Moisan, J., O'Reilly, J.E., Haidvogel, D., 2006. Nitrogen cycling in the Middle Atlantic Bight: Results from a three-dimensional model and implications for the North Atlantic nitrogen budget. Glob. Biogeochem. Cycles 20, 14. doi:10.1029/2005GB00245.

Schartau, M., Oschlies, A., 2003. Simultaneous data-based optimization of a 1D-ecosystem model at three locations in the North Atlantic: Part I - Method and parameter estimates. J. Mar. Res. 61, 765–793.

Moore, J.K., Doney, S., Lindsay, K., 2004. Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. Glob. Biogeochem. Cycles 18(4).

Another parameter that appears rather low is the maximum growth rate of the au- totrophic diazotrophs. For example, Holl & Montoya, J. Phycol. 44:929 (2008) re- ported growth rates greater than 0.6/d for Trichodesmium grown in a chemostat, so a maximum (actually potential) rate parameter of 0.25/d appears unrealistically low. My impression is that these low settings reduce diazotrophy too much, maybe just compensating for the assumption of obligate diazotrophy but maybe also being responsible for the requirement of aphotic N2 fixation in the present model.

Response: Please note that the maximum growth rate in the model is temperature dependent and that 0.25/d is the reference value at 0°C (as stated in Table 3). At typical water temperatures in the Gulf of Aqaba of 20°C this results in an actual maximum growth rate of 0.97/d, close to the value mentioned by the Reviewer.

Further, the authors say that the diazotroph parameters were unconstrained by the data and that the parameter setting were taken from the literature, but do not provide references in Table 3 or elsewhere. The ms also does not say how it was determined that the parameters were unconstrained by the data. This seems inappropriate to me, since this is specifically a model study about diazotrophy, so I expect that great care is taken to select appropriate parameter settings. The fact that the diazotroph parameters are unconstrained by the data makes the choice of data appear questionable to me. In my view, the data should be able to constrain the most important aspects of a model's performance, and if this is not the case, one should try to either find better data or develop a better cost function (see below). The problem is that the inability to constrain the model parameters with the data implies that the associated processes are actually irrelevant. The simple fact that the authors observe better model performance when including diazotrophs implies that the associated parameters must have an effect, so I expect that a better cost function can in fact be designed which is capable of constraining those paramenters.

Response: With regard to the Reviewer's assertion that "parameter setting were taken from the literature, but do not provide references in Table 3 or elsewhere," we would like to point to Table 2 which lists ranges for each parameter from the published literature with the corresponding references (see columns 3 and 6). Table 3 now emphasizes these references too, only included in the text previously.

In response to the comment that we do "not say how it was determined that the parameters were unconstrained by the data": Showing that measurements of chlorophyll, nutrients and oxygen do not constrain N_2 fixation rates is not within the intended scope of this paper. In fact, it appears self-evident to us. What we here mean by "unconstrained" is that, for example, chlorophyll alone cannot provide any information about how much chlorophyll is due to diazotrophs and how much is due to non-diazotrophs, therefore this variable alone is unable to help in the determination of diazotrophic parameters. The systematic calibration method we use relies on having direct observational counterparts (i.e., from the same location at least) to the model output. From previous knowledge and experience using optimization methods, we would need comprehensive N_2 fixation and/or size-structured diazotrophic biomass data in order to constrain diazotrophs parameters.

With regard to the comment that "the choice of data appear questionable to" the Reviewer, we would like to respond that we used all the data that was available. The suite of available measurements is very typical of multi-year oceanographic time series (most aren't as comprehensive and well-funded as the HOT and BATS programs).

With regard to the cost function we refer to the last paragraph in our response to the first comment.

3. Model evaluation. The authors report that they performed sensitivity analyses to obtain information of sensitive model parameters but they do not say how the sensitivity was quantified nor present any results from the sensitivity analyses. This could well be done in the supplement, but it is important for those who want to work with the model later.

Response: The information about the preliminary sensitivity analysis in now included in the Supplement.

The authors mention that they considered the first year of the model simulations as spinup but do not say how the model was initialised (from observations? what about the non-observed variables?). From my own experience with 1D modelling, one year is a rather short period for a spinup. Did the authors try longer spinups in order to find out whether the model is sufficiently close to a quasi-steady-state after one year? This should be discussed as well. It is this kind of omission, together with missing entries in the list of references (e.g., Fernandez 2011 and Smith 1936), that leaves an impression of sloppiness.

Response: We have included this information in the methods section as follows:

"*NO3, NH4 and PO4 initial total concentrations match the observed total inventories, using homogenous concentrations throughout the water column of 2.5 mmol m-3, 0.05 mmol m^{-3} and 0.15 mmol m^{-3}, respectively. Vertical nutrient concentrations are redistributed within few months and replicate the observed vertical distributions well starting from October 2005. Non-fixing phytoplankton, zooplankton and detritus are initially set to a homogeneous small value of 0.1 mmol N m^{-3}, which also readjust rapidly because the adjustment timescales for these variables are short (Fennel et al., 2006). Diazotrophs initial values are set to lower densities than non-fixing organisms, with a homogenous total value of 0.03 mmol N m^{-3} (i.e., models with multiple diazotrophs maintain the same amount of initial diazotrophic biomass).*"

4. Parameter estimation. The authors apply RMSEs of absolute concentrations to obtain a measure of model-data misfit. This cost function will not be sensitive to large relative deviations if the absolute concentrations are low. Thus, it is only logical that the inability of the model to reproduce the negative N* in the surface waters "is not a source of large data-model discrepancies" (l. 8, p. 12). Introducing relative-error information or local scaling into the cost function could help here. The most important shortcoming of the authors' cost function, however, is that it neglects error

correlations, see, e.g., Schartau et al., Biogeosci. 14:1647 (2017).

Response: Our cost function does indeed scale depending on the data type by weighing the contributions from the different variables by their standard deviations.

5. Figures. The use of log-scales in Fig. 9 makes it impossible to see the differences among models and between models and data. Please use a linear scale.

Response: We chose the log-scale to be able to visualize differences at very low values, common for these rates in the region.

6. Conclusions. As it stands, the conclusions are not sufficiently supported be the model analysis described. In particular, the conclusions about aphotic N2 fixation are compromised by the choice of unrealistic parameter values constraining autotrophic diazotrophy to the very surface. If inferences about heterotrophic diazotrophy are to be drawn, at least the parameters determining the depth distribution of autotrophic diazotrophy must be analysed with a detailed sensitivity analysis. The current analysis cannot say whether the deep N signal is really due to aphotic N2 fixation or exported material from the surface.

Response: As per our responses above, the parameters for autotrophic diazotrophs are fully consistent with the existing literature.

2. SUMMARIZED LIST OF CHANGES

- We changed our manuscript title to better reflect the scope of our results.
- We changed our focus on N* and removed plots and discussion concerning this tracer, following reviewers suggestions.
- We performed an additional model experiment (H3'), which served as a control of the effect of adding a non-fixing heterotrophic organism. Vertical distributions of nitrate, phosphate, chlorophyll and oxygen results are shown. Our main study conclusions were not affected.
- We included a sensitivity analysis to changes in the N:P constant ratios of N_2 fixing and non-fixing organisms. Our main study conclusions were not affected.
- We extended the discussion, in particular the section concerning limitations and uncertainties. This was done to include relevant studies suggested by our reviewers.

- We extended the supplement to include a more detailed description of the equations of all model versions.
- We extended the supplement to include a description of the preliminary parameter sensitivity analysis cited in the text.
- We clarified relevant details in our methods, such as how denitrification is implemented in the models.
- We changed the use of "mechanistic models" to "trait-based models."
- We specified references followed for diazotrophic parameters in the corresponding table. Previously these were cited only in the text.
- We changed the colorscale of results plots to a colorblind friendly colormap.

3. MARKED-UP MANUSCRIPT VERSION

In yellow sections that have been modified.

Modelling the biogeochemical effects of heterotrophic and autotrophic N_2 fixation in the Gulf of Aqaba (Israel), Red Sea

Angela M. Kuhn[1, 2], Katja Fennel[1], Ilana Berman-Frank[3, 4]

[1]Department of Oceanography, Dalhousie University, Halifax, B3H 4R2, Canada
[2]Scripps Institution of Oceanography, University of California San Diego, La Jolla, 92093-0021, USA*
[3]Mina and Everard Goodman Faculty of Life Sciences, Bar Ilan University, Ramat Gan, 5290002, Israel
[4]Leon H. Charney School of Marine Sciences, University of Haifa, Mt. Carmel, Haifa, 3498838, Israel*

*current address

Correspondence to: Angela M. Kuhn (angela.kuhn@dal.ca)

Abstract. Recent studies demonstrate that marine N_2 fixation can be carried out without light by heterotrophic N_2-fixers (diazotrophs). However, direct measurements of N_2 fixation in aphotic environments are relatively scarce. Heterotrophic, as well as unicellular and colonial photoautotrophic diazotrophs, are present in the oligotrophic Gulf of Aqaba (northern Red Sea). This study evaluates the relative importance of these different diazotrophs by combining biogeochemical models with time series measurements at a 700m-deep monitoring station in the Gulf of Aqaba. At this location, an excess of nitrate, relative to phosphate, is present throughout most of the water column and especially in deep waters during stratified conditions. A relative excess of phosphate occurs only at the water surface during nutrient-starved conditions in summer.

We show that a model without N_2 fixation can replicate the observed surface chlorophyll but fails to accurately simulate inorganic nutrient concentrations throughout the water column. Models with N_2 fixation improve simulated deep nitrate by enriching sinking organic matter in nitrogen, suggesting that N_2 fixation is necessary to explain the observations. The observed vertical structure of nutrient ratios and oxygen is reproduced best with a model that includes heterotrophic, colonial and unicellular autotrophic diazotrophs. These results suggest that heterotrophic N_2 fixation contributes to the observed excess nitrogen in deep water at this location. If heterotrophic diazotrophs are generally present in oligotrophic ocean regions, their consideration would increase current estimates of global N_2 fixation and may require explicit representation in large-scale models.

1 Introduction

Biological nitrogen fixation refers to the conversion of dinitrogen gas (N_2) via reduction to ammonium (NH_4) into bioavailable forms of nitrogen by a specialized group of microbes containing the nitrogenase enzyme complex. On geological timescales, the size of the oceanic reservoir of bioavailable nitrogen, and thus the ocean's capacity for exporting carbon to depth, is controlled by the balance between removal of fixed nitrogen by denitrification and input by N_2 fixation (Falkowski, 1997; Haug et al., 1998; Deutsch et al., 2007; Gruber and Galloway, 2008; Fennel et al., 2005). The amount of organic matter exported from the surface to the deep ocean (i.e., export production) depends on allochthonous inputs of nitrogen (i.e., "new nitrogen") into the euphotic zone (Eppley and Peterson, 1979). These new nitrogen inputs determine the amount of "new production", which is directly related to the exported fraction. Locally the supply of new nitrogen can occur through several mechanisms, including microbially mediated N_2 fixation, diapycnal mixing injecting deep nitrate (NO_3) into the surface, lateral transport, atmospheric sources and riverine input. While the injection of deep NO_3 is often regarded as the dominant source of new nitrogen that drives the seasonal cycle of marine primary production; there is significant interest in quantifying the contribution of N_2 fixation to primary production, particularly in oligotrophic areas (Karl, 2002; Zehr and Ward, 2002; Capone et al., 2005; Luo et al., 2012).

Diazotrophs contain nitrogenase, the catalyst enzyme for N_2 reduction, which is encoded by *nif* genes. *Trichodesmium spp.,* a group of non-heterocystous filamentous cyanobacteria that form large colonies, were traditionally considered the main contributors to N_2 fixation in the surface subtropical and tropical ocean (Carpenter and McCarthy, 1975; Capone et al., 2005). Increased

sampling efforts and method improvements subsequently led to the discovery of diverse diazotroph groups including heterocystous endosymbiotic cyanobacteria (Zehr et al., 1998; Carpenter et al., 1999), free-living unicellular cyanobacteria (Zehr et al., 2001; Montoya, 2004; Moisander et al., 2010), and other cyanobacterial symbionts (Zehr et al., 2000). Most recently, genetic techniques have allowed the detection of *nif* genes in a number of anaerobic and heterotrophic phylotypes (Zehr et al., 2008; Zehr, 2011; Rahav et al., 2013, 2015). The abundance of *nif* genes does not necessarily imply that these organisms are actively fixing N_2 (Zehr et al., 2000; Moisander et al., 2017); however, the correlation between bacterial productivity and N_2 fixation rates suggests that significant aphotic N_2 fixation may occur in the Red Sea (Rahav et al., 2013, 2015). The differences in size and physiology of these diverse diazotrophs also suggest that they occupy distinct niches, and thus may affect primary productivity and export production differently (Bonnet et al., 2016; Moisander et al., 2010).

Most biogeochemical models treat N_2 fixation as a purely light-dependent, autotrophic process. These models either use standard formulations of light limitation for the diazotrophic groups (e.g., Fennel et al., 2002; Moore, et al., 2004; Gregg, 2008; Dutkiewicz et al., 2012), or include theoretical considerations to introduce an empirical N_2 fixation flux in the model (e.g., Bisset et al., 1999). Some approaches neglect light limitation on diazotrophy and instead infer global N_2 fixation patterns from the distribution of dissolved inorganic nitrogen and phosphorus, and estimates of ocean circulation (e.g., Deutsch et al., 2007). In trait-based models, diazotrophs are usually accounted for by a single functional group with biological rate parameters intended to represent either *Trichodesmium spp.*, unicellular cyanobacteria, or a generic autotrophic diazotroph. More complex models of diazotroph growth have explored cellular function and the effects of a variable cellular stoichiometry, however still limited by light (Kreuss et al., 2015; Fernández-Castro et al., 2016). Only a few modelling studies have evaluated multiple autotrophic diazotroph groups simultaneously, by considering separate groups for *Trichodesmium spp.*, unicellular cyanobacteria and diatoms-cyanobacterial associations (e.g., Monteiro et al., 2010; Dutkiewicz et al., 2012, 2015). To our knowledge, heterotrophic N_2 fixation has not yet been considered explicitly in biogeochemical models.

Understanding the ecological dynamics of different types of diazotrophs should significantly improve predictive capabilities in biogeochemical models, and lead to more accurate estimates of global N_2 fixation rates. It has been suggested that N_2 fixation rates are underestimated globally due to limited knowledge about the distribution and characteristics of N_2 fixing organisms (Montoya, 2004; Zehr, 2011). It is also assumed that marine N_2 fixation may increase globally, as

a result of ocean warming and higher concentrations of dissolved CO_2 in sea water (Hutchins et al., 2007; Levitan et al., 2007; Dutkiewicz et al., 2015). To our knowledge, these laboratory experiments have only explored the response of *Trichodesmium*. Less information is available about the effects of climate trends on other diazotrophic organisms.

In this study we explore the biogeochemical signatures that result from different assumptions about the ecological niches occupied by diazotrophs. Our study area is the Gulf of Aqaba, a northern extension of the Red Sea. Aside from the reported presence of diverse diazotroph types, the morphology of the Gulf of Aqaba limits horizontal transport of deep waters, thus allowing us to simplify the physical model's complexity and focus on the biological component. We aim to answer the following two questions: i) How important is N_2 fixation as a source of new nitrogen in the Gulf of Aqaba? ii) How important is heterotrophic, light-independent N_2 fixation? To address these questions, we implemented a one-dimensional model at a monitoring station for which monthly quality-controlled measurements of physical and biogeochemical variables are available from 2004 onward. We then systematically tested different model assumptions about diazotrophy and calibrated selected model parameters to facilitate an objective comparison between the different biogeochemical model versions. The different assumptions about diazotrophy consider the characteristics of organisms identified in the Gulf of Aqaba, including heterotrophic fixing α and γ proteobacteria (Rahav et al., 2013), unicellular cyanobacteria, and *Trichodesmium spp.* (Post et al., 2002; Foster et al., 2009; Rubin et al., 2012). Our most important conclusion is that aphotic N_2 fixation is necessary to reproduce the observed excess nitrogen in deep waters of the Gulf, while maintaining reasonable surface N_2 fixation rates. Our best performing biogeochemical model of the Gulf of Aqaba estimates annual N_2 fixation rates in overall agreement with local and large-scale estimates in the literature.

2 Study Area: The Gulf of Aqaba

The Gulf of Aqaba is a quasi-rectangular, 200-km long, 20-km wide, semi-enclosed basin in the northeast region of the Red Sea (Figure 1). The average depth of the Gulf of Aqaba is 800 m, its deepest point approximately 1800 m, and it is surrounded by arid mountains that steer the dominantly northerly winds (Berman et al., 2003). Two shallow sills, the Bab el Mandeb (~140 m) and the Strait of Tiran (~240 m), inhibit the entrance of cold and dense deep waters from the Indian Ocean. Since inflow is restricted to warm surface waters, the Gulf's deep-water masses (>300 m) are locally formed (Wolf-Vecht et al., 1992; Biton et al., 2008) and have negligible horizontal transport toward the exterior (Klinker et al., 1976; Manasrah et al., 2006).

The annual hydrographic cycle exhibits a well-defined seasonality where vertical temperature and salinity distributions are dominantly affected by surface heat fluxes and modified by surface advective fluxes (Carlson et al., 2014). During winter (December to March), convective vertical mixing usually extends to depths >300 m (Labiosa et al., 2003), and even reaches the bottom (~700-800 m bottom depth) in some extreme years (Figure 2). From April to September, the water column is thermally stratified, and inflowing warm surface waters from outside the Gulf occupy the layer above the thermocline (Genin and Paldor, 1998; Berman et al., 2000; Biton and Gildor 2011). During fall (October to December), surface cooling and high evaporation rates erode the seasonal stratification and re-establish a well-mixed water column (Berman et al., 2003; Monismith and Genin, 2004). The Gulf experiences net evaporation of approximately 1.6 m yr^{-1} (Ben-Sasson et al., 2009) due to negligible precipitation and run-off (Wolf-Vecht et al., 1992).

The Gulf is oligotrophic, with surface nitrate (NO_3) and phosphate (PO_4) concentrations usually close to their detection limits during summer stratification (Fuller et al., 2005; Mackey et al., 2009; Meeder et al., 2012). Deep winter mixing supplies inorganic nutrients to the surface, and NO_3 and PO_4 reach ~0.1 μM and ~2 μM, respectively (Figure 2; Lindell and Post 1995; Lazar et al. 2008). Dust from the desert provides a sufficient atmospheric source of soluble iron (Fe) for microbial growth in the Gulf (Chase et al., 2006; Chen et al., 2007). Phytoplankton spring blooms reach a maximum chlorophyll concentration of ~2 mg Chl-a m^{-3} and their initiation is strongly correlated with the termination of winter cooling of the water column (Zarubin et al., 2017). Interannual variability in the depth of winter convective mixing results in periods of nutrient accumulation in deep waters (Figure 2; Wolf-Vecht et al., 1992; Lazar et al., 2008; Carlson et al., 2012), which is re-set during extreme winter mixing events approximately every four years (Silverman and Gildor, 2008). The periodicity of these extreme mixing events has been associated with regional weather patterns that modify the Red Sea water temperatures (Silverman and Gildor, 2008).

3 Methods

We analysed the role of autotrophic and heterotrophic N_2-fixing organisms in determining biogeochemical patterns at an open pelagic site (Station A), located in the northern Gulf of Aqaba, by testing four alternative ecosystem model versions. The ecosystem models are evaluated in terms of their ability to replicate observations of oxygen (O_2), NO_3, PO_4, and chlorophyll. In this section, we first describe the available observations, then the models, and finally the systematic model calibration method.

3.1 Observations

Meteorological and oceanographic observations are available from the Inter-University Institute (IUI) for Marine Sciences in Eilat, Israel (http://www.iui-eilat.ac.il/Research/NMPMeteoData.aspx). Meteorological observations are used to calculate surface heat and momentum fluxes for the physical model and incoming light for the biological models. Observed meteorological variables include wind speed, air temperature, air humidity, air pressure, irradiance and cloud cover. This data has been collected continuously and automatically at 10 min intervals by the meteorological instrumental array at the end of the IUI pier since 2006.

Monthly CTD and bio-chemical profiles at Station A (29.5° N, 34.9° E) were collected during monthly surveys of the National Monitoring Program (NMP) from 2004 to 2014 (http://www.iui-eilat.ac.il/Research/NMPMeteoData.aspx). CTD profiles are used to nudge temperature and salinity in the physical model (see section 3.2). Bio-chemical profiles, including NO_3, nitrite (NO_2), ammonium (NH_4), PO_4, O_2 and chlorophyll-a (Chl-a), are used for biogeochemical model calibration (years 2006 to 2010) and model validation against unassimilated data (years 2011 to 2014). Nutrients were measured using spectrophotometry (QuickChem 8000 flow injection), O_2 was determined by Winkler titrations, and Chl-a concentrations are estimated using fluorometry (Turner Designs 10-AU).

3.2 Model Descriptions

The ecosystem models are implemented within the General Ocean Turbulence Model (GOTM), a one-dimensional physical model that computes solutions to differential equations for the vertical transport of momentum, salt and heat (Burchard et al., 1999). GOTM is implemented for the 700-m deep station with a vertical resolution of 3 m and forced with hourly meteorological observations from the IUI pier. Temperature and salinity are nudged to observed CTD profiles with a nudging time scale of 30 days. This is done to account for the influence of horizontal advection of heat and salt in the one-dimensional model and ensures a realistic representation of density stratification. The effect of temperature and salinity nudging on the results is analysed below (section 4.2.1). As model calibration is computationally expensive, model simulations run only from January 2005 to September 2010. The first year of each simulation is considered model spin-up and excluded from further analysis (climatological meteorological forcing is used for the first year, as this database starts in 2006). NO_3, NH_4, PO_4 and O_2 initial total concentrations match the observed total inventories, using homogenous concentrations throughout the water column of 2.5 mmol m^{-3}, 0.05 mmol m^{-3} and 0.15 mmol m^{-3}, respectively. Vertical nutrient and

oxygen concentrations are redistributed within a few months and replicate the observed vertical distributions well starting from October 2005. Non-fixing phytoplankton, zooplankton and detritus are initially set to a homogeneous small value of 0.1 mmol N m^{-3}, and also readjust rapidly because the adjustment timescales for these variables are short (Fennel et al., 2006). Diazotrophs initial values are set to lower densities than non-fixing organisms, with a homogenous total value of 0.03 mmol N m^{-3} (i.e., models with multiple diazotrophs maintain the same amount of initial diazotrophic biomass by dividing initial values equally among organisms).

Four main ecosystem model versions of increasing complexity (referred to as H0, H1, H2 and H3) are treated as alternative hypotheses of how biological processes, especially diazotrophy, control the vertical distribution and temporal variability of dissolved inorganic nutrients and oxygen. H0 is the base model without diazotrophics and follows the model equations described in Fennel et al., (2006, 2013). In general, the model follows Monod kinetics using a fixed N:P ratio ($R_{N:P}^{nf}$= 16). Sensitivity to the constant N:P ratio is explored in section 4.2.3. We test the H0 model with and without a sediment denitrification flux (model versions H0 and H0', respectively). H0 includes denitrification but no N_2 fixation, as no diazotrophs are considered. H0' does include neither denitrification, nor diazotrophic organisms, and thus the underlying assumption of this model version is that there is a balance between inputs from N_2 fixation and losses of fixed nitrogen due to denitrification. When present, the denitrification flux follows Fennel et al. (2013) with a loss fraction of 6 mol N_2 per mol of organic matter remineralized at the sediment-water interface. This generates an average sediment denitrification flux of 0.25 ± 0.46 mmol N m^{-2} d^{-1}, with a maximum value of 3.01 mmol N m^{-2} d^{-1}.

H1, H2 and H3 are modified versions of H0, in which different groups of diazotrophic organisms are added sequentially. Diazotrophic growth formulations are similar to those of non-fixing phytoplankton, except that they are not limited by nitrogen and have a higher N:P ratio ($R_{N:P}^{f}$ = 45). H1 introduces a generic autotrophic diazotroph; H2 replaces H1's generic diazotroph with two autotrophs representing unicellular and colonial (e.g., *Trichodesmium spp.*) cyanobacteria. The unicellular group overall follows the same formulation as the generic diazotrophs, except that no aggregation term is included. We simplify unicellular diazotroph behaviour in this manner because this group represents free-living picoplanktonic cells that typically do not form large colonies although they can aggregate (Bonnet et al. 2016). Instead, we assume this group is grazed by zooplankton at similar rates as non-fixing phytoplankton. This difference between colonial and unicellular groups is consistent with studies suggesting that colonies represent an evolutionary adaptation to decrease grazing pressure (Nielsen, 2006). Aside from their size,

Trichodesmium spp. colonies may be less palatable and harder to digest due to toxins (Kerbrat et al. 2010, 2011). Grazing is not a major fate of this group (O'Neil and Roman, 1994).

The last model version, H3, adds a heterotrophic group to the model structure of H2. This functional group is not limited by light and grows by consuming both dissolved inorganic and organic forms of nutrients. An intermediate version H3' is used as a control, where the heterotrophic organisms do not fix nitrogen and are limited by the availability of nitrogen in inorganic forms and from small detritus. Model H3 eliminates the nitrogen limitation and the heterotrophic group becomes a heterotrophic diazotroph group. In a subsequent set of four experiments (H3a, H3b, H3c, and H3d), we remove complexity from H3. The heterotrophic group remains, but we sequentially remove the autotrophic groups one-at-a-time: first the colonial cyanobacteria (H3a), then the unicellular cyanobacteria (H3b), then the generic autotrophic diazotroph (H3c), and finally we remove all autotrophic diazotrophs (H3d). A summary of all model versions is given in Table 1 and a description of state variables and full model equations in the Supplement.

We acknowledge that some of these model assumptions are still simplifications of diazotroph behaviour. For example, *Trichodesmium spp.*, as well as other autotrophic and diazotrophic organisms, may also take up dissolved organic matter to support growth (Benavides et al., 2017). Nevertheless, model assumptions and formulations used here are in line with the most commonly accepted understanding of the dominant controls on microbial and diazotrophic growth.

3.3 Model Parameters

3.3.1 Parameter Optimization Method

Parameter optimization refers to the minimization of misfit between model and observations by adjusting model parameters. We applied the method first to systematically calibrate the most sensitive parameters of H0 (see supplement), and then to independently re-calibrate parameters in H1 to H3 after the introduction of diazotrophs. We used an evolutionary algorithm, where changes in the parameter values follow a set of rules inspired by the process of natural selection (Houck et al., 1995; Kuhn et al., 2015). The algorithm starts with a randomly generated "population" of 30 parameter sets ($\vec{p}$), which are iteratively modified over a number of generations. During each generation of the population, the cost $J(\vec{p})$ of the model with parameter set $\vec{p}$ is calculated as:

$$J(\vec{p}) = \frac{1}{V}\sum_{v=1}^{V}\frac{w_v}{N}\sum_{i=1}^{N}\left(\hat{y}_{v,i} - y_{v,i}\right)^2, \quad (1)$$

where $\hat{y}$ represents a model estimate and y the corresponding observation. *N* is the number of observations included for each variable v. Here the number of variables *V* is 5 (nitrate + nitrite, ammonium, phosphate, chlorophyll-a, and oxygen measured as profiles at Station A between 2006 and 2010). Model-data misfits are weighted by the factor $w_v = 1/\sigma_v$, i.e, the inverse standard deviation of each variable. Half of the parameter sets with the lowest *J* value "survive" to the next generation. The other half of the population is regenerated from new parameter sets obtained by recombination of two random "parent" sets drawn from the better performing half (i.e., the "survivors" of the previous generation). Parameters also "mutate", i.e. random noise is added, for additional variability in the parameter space. An allowable range of values is set for each parameter based on the literature (Table 2).

3.3.1 Optimized Parameters

The parameter optimization method has limitations. Most importantly, the optimization cannot estimate with confidence parameters that are unconstrained by the observations (Fennel et al. 2001; Schartau and Oschlies, 2003; Ward et al., 2010). To avoid this, a subset of H0's most sensitive parameters was selected for optimization through a preliminary sensitivity analysis. Optimized parameters for H0 are identified in Table 2 along with the optimal values. The optimization was replicated 10 times over 100 generations using the algorithm described in section 3.3.1. Non-optimized parameters are fixed at their a priori estimates based on Fennel et al. (2006, 2013).

For each model version with diazotrophs (H1, H2 and H3), some of the parameters already optimized for H0 required re-calibration to properly accommodate the changes in system dynamics. Re-calibrated parameters for each model version are presented in Table 3. No re-calibration was performed for model versions (H0' and H3a-d), as they are aimed to test the relative importance of individual model components.

3.3.2 Diazotroph Parameters

Since none of the parameters directly related to the diazotroph groups are constrained by the available observations, they were predefined for H1, H2 and H3, based on the observational and modelling literature (Table 3). Previous modelling studies have used maximum growth rates of

generic N_2 fixers ranging from 0.4 d^{-1} (Moore, et al., 2004) to 1.25 d^{-1} (Ward et al., 2013). When model diazotrophs are assumed to represent *Trichodesmium spp*. values range between 0.17 d^{-1} (Hood et al., 2001) to 0.3 d^{-1} (Fennel et al., 2002). From the observational literature, *Cyanothece* (unicellular cyanobacteria) and *Trichodesmium spp.* cultured under various combinations of Fe and light availability exhibit maximum rates around 0.3 ± 0.05 d^{-1} (Capone et al., 1997; Berman-Frank et al., 2001; Hutchins et al., 2007). Growth rates can be higher (up to 0.5 d^{-1}) at high CO_2 and high light availability (Kranz et al., 2010; Hong et al., 2017). We chose a common reference maximum growth rate of 0.25 d^{-1} for all photosynthetic diazotrophs, such that differences between the model versions result only from the different assumptions about the losses of each group (e.g., predation of unicellular cyanobacteria vs. sinking of large aggregates). Based on growth rates measured for cultured heterotrophic bacteria, we chose a value of 0.2 d^{-1} for the heterotrophic diazotrophs (Pomeroy and Wiebe, 2001). Observational and modelling studies were also considered to set the photosynthetic initial slope of photosynthetic diazotrophs (Geider et al., 1997; Moore et al., 2004; Hutchins et al., 2007). Other parameters are based on Fennel et al. (2002).

4 Results

4.1 Observed NO_3 and PO_4 Patterns

To provide context for the evaluation of our model simulations, we first describe the observed interannual and seasonal variability of NO_3 and PO_4 for the complete time series (2004 to 2014) at Station A (Figure 2). From May to January vertical distributions of NO_3 and PO_4 show depletion of nutrients in the euphotic zone and a nutricline between 100 and 200 m. From February to April, nutrient concentrations increase near the surface and decrease in deep waters (>200 m) as result of vertical mixing. Multi-year periods of accumulation of nutrients in deep waters were observed from: i) the beginning of the series to the end of 2006, ii) after the winter of 2008 until February 2012, and iii) after the winter of 2013 until the end of the series. These periods are bookended by winters with extremely deep mixing events in 2007, 2008, 2012 and 2013 during which nutrient concentrations are nearly homogenized in the entire water column. Two prolonged periods of these vertically homogenous conditions were observed in 2007 and 2008, lasting two to three months. Our model calibration simulations are from 2006 to 2010, allowing us to include two years with deep winter mixing (2007 and 2008) and two years with moderate winter mixing (2009 and 2010). (deleted paragraph)

4.2 Model Results

4.2.1 Sensitivity to Physical Nudging

Model runs, with and without temperature and salinity nudging towards observations, demonstrate that nudging has a negligible effect below 200 m, indicating that horizontal advection does not modify the lower part of the water column in a significant way (Supplement). Above 200 m, nudging corrected model errors in the representation of vertical mixing and surface forcing. The average magnitude of the differences due to nudging in the top 200 m is 0.20±0.45 ºC and 0.5±0.16 kg m^{-3}. Since these effects are small and limited to the surface, we conclude that neither nudging nor the neglect of horizontal advection affects our conclusions significantly.

4.2.2 Effects of N_2 Fixation on DIP and DIN

Figure 3 shows simulated NO_3 and PO_4 concentrations from models H0', H0, H1, H2, H3' and H3, along with the corresponding measurements. Observed NO_3 and PO_4 concentrations exhibit a marked increase in deep water after the strong winter mixing of 2008. Weaker winter mixing after 2008 results in deep-nutrient accumulation, which is more pronounced for NO_3 than PO_4. Model H0' reproduces some deep-nitrate accumulation, but underestimates NO_3 concentrations in comparison to observations. Model H0 strongly underestimates inorganic nitrogen below the nutricline. Model H1, where N_2 fixation was introduced via a generic autotroph, generates only small changes in the vertical distribution of nutrients. In model H2 the representation of NO_3 below the nutricline is slightly improved; however, underestimation of mid-water NO_3 is still noticeable. Model H3 significantly improves the representation of deep NO_3 accumulation. All model versions represent similar vertical distributions of PO_4 and underestimate its deep-water concentrations by the end of the series.

These model differences are also summarized in Figure 4, which shows the simulated and observed NO_3 and PO_4 inventories in surface and deep waters. According to the observations deep NO_3 accumulates between 2007 and 2010 at a rate of 0.59 ± 0.08 mmol m^{-2} d^{-1}, whereas deep PO_4 accumulates at 0.015 ± 0.009 mmol m^{-2} d^{-1}. During this accumulation period approximately 36 mmol NO_3 per mmol PO_4 appear in deep waters. All five model versions simulate similar magnitudes and temporal variability of PO_4, but NO_3, in particular below 100 m, diverges over time among the models. H0 has the largest deviations from the other models, simulating approximately constant deep NO_3 after 2007. H0', the version without denitrification, produces a rate of increase in deep NO_3 similar to that of model version H2. H3 has the highest accumulation rate of deep NO_3, matching the observed slope the best.

Of the additional model versions based on H3 (H3a, H3b, H3c and H3d), results from H3a (heterotrophic and colonial diazotrophs) come closest to H3, while the model without autotrophic diazotrops has the narrowest and most unrealistic range of nutrient concentrations (not shown).

4.2.3 Sensitivity to Planktonic Stoichiometry

To investigate whether the vertical distribution of dissolved inorganic nutrients was affected by our assumption of fixed N:P phytoplankton and diazotroph ratios, we explored the sensitivity of NO_3 and PO_4 to changes in the non-fixing phytoplankton N:P ratio ($R^{nf}_{N:P}$= 16) and the diazotrophs N:P ratio ($R^{f}_{N:P}$= 45). In this analysis we used model version H1, which includes a single non-fixing group and a single N_2 fixing group and varied the ratios one-at-a-time. The range of values for $R^{nf}_{N:P}$ varied from 10 to 28 and the range for $R^{f}_{N:P}$ from 19 to 51. Figure 5 shows examples of the results obtained by increasing and decreasing each ratio.

Changes in the N:P ratios had negligible effects on the vertical distribution of NO_3, but strongly affected PO_4 distribution. In general, lower than Redfield $R^{f}_{N:P}$ of diazotrophs increases PO_4 below 300 m, most strongly at depth. This occurs possibly because more phosphorus returns to the dissolved pool per unit nitrogen trough excretion and remineralization. In contrast, a decrease of more than half the $R^{nf}_{N:P}$ of non-diazotrophs produces only a minor decline in deep PO_4, while increases in the ratio did not have a significant effect. The decline in deep PO_4 occurs because, in the absence of nitrogen limitation, diazotrophs can utilize additional phosphorus.

4.2.4 Effect of N_2 Fixation on Chlorophyll and O_2

Figure 6 shows simulated and observed chlorophyll and dissolved oxygen values. The seasonal variability of total chlorophyll concentrations is reproduced well by all models, with higher chlorophyll between November and April. During these months, simulated chlorophyll concentrations are homogeneous down to 200 m. In 2007 and 2008, chlorophyll concentrations of ~0.13 mg m^{-3} are observed in the measurements reaching as deep as 500 m. This feature is also captured well by our models, as is the location of the deep chlorophyll maximum (DCM) at ~80 m between March and October. However, there are some discrepancies between model results and observations. The models overestimate spring bloom peak concentrations in 2007 and predict peak timing two months earlier than observed in 2008. Model H0 tends to underestimate chlorophyll concentrations from the surface to the DCM during summer months. As chlorophyll concentrations are extremely low during this time of the year, these model-data differences are on the order of 0.05 to 0.1 mg m^{-3}. These discrepancies during summer months are corrected in the models with N_2 fixation.

Simulated oxygen concentrations exhibit larger differences between models and observations, in particular below the mixed layer, where air-sea fluxes do not directly affect oxygen concentrations. Model versions without diazotrophs (H0 and H0') show similar deep-oxygen variability, with a small underestimation of oxygen during the winter of 2007, and a small overestimation after the winter of 2008. Model re-calibration for the model versions with diazotrophs results in changes in deep oxygen. H1, the model with generic diazotrophs, exhibits the largest model-observation misfits. As in the case of deep NO_3, the best deep-oxygen representation is obtained with H3.

4.2.5 Validation against Independent Observations

Observations from 2010 to 2014 (outside the optimization period) are used to independently validate the models. The root-mean-square errors (RMSEs) in Table 4 show that, in terms of chlorophyll, PO_4 and surface O_2, all models behave similarly and achieve similar agreement for assimilated and independent observations. As demonstrated in the previous sections, the model versions mainly diverge in their behaviour with respect to NO_3, with some differences in O_2 concentrations. Between 0 and 100 m, H3 has the largest RMSEs for NO_3, but below 100 m it has the lowest values, particularly against unassimilated NO_3. H3 also has the lowest RMSEs for surface and deep oxygen (Table 4).

Figure 7 shows observed and simulated NO_3 inventories in 0 – 100 m and below 100 m outside the assimilation period. Compared against the other model versions, H3 increasingly overestimates surface NO_3 over time. However, the deep NO_3 inventory is best represented by H3. By the end of the observed time series, between 2013 and 2014, H3 starts to also overestimate deep NO_3.

4.2.6 Primary Production and N_2 Fixation Rates

We now compare the simulated rates of primary production with those reported for the Gulf of Aqaba by Rahav et al. (2015) and Iluz et al. (2009) (Figure 8a) and the simulated rates of N_2 fixation with those measured by Rahav et al. (2015) and Foster et al., (2009) (Figure 8b). Following Rahav et al. (2015), we show the rates at the DCM and their averages above and below the DCM. The depth-resolved in situ primary production rates reported by Iluz et al. (2009) were also averaged in the same way for comparison. Where necessary, observations of primary production in carbon units were converted to the model's nitrogen units using the Redfield ratio.

Simulated primary production above the DCM ranges from 0.02 to 0.85 mmol N m^{-3} d^{-1} and exhibits an annual cycle with peaks of productivity in October and April. A prolonged period of

low primary production extends from April to September in most model versions. Model versions H3b and H3d maintain rates twice as large as the rest of the models during the summer/fall period. Aside from H3b and H3d, differences between models are small and simulated rates agree with those measured by Iluz et al. (2009) and Rahav et al. (2015).

Above the DCM, models H1, H2, H3 and H3a show a well-defined N_2 fixation peak during summer months (i.e., after the peak in primary production). Maximum rates in these models range from 0.001 to 0.1 mmol N m^{-3} d^{-1}, which agrees with the observed rates by Foster et al. (2009) and Rahav et al. (2015). In general, simulated N_2 fixation rates are low during winter and spring. Similar temporal patterns and differences between model versions occur at the DCM and below. Peaks in N_2 fixation at these depth levels occur after the surface peak, and have a shorter duration and smaller amplitude. Deep N_2 fixation rates estimated by models without heterotrophic diazotrophs do not match the observed rates by Rahav et al. (2015).

5 Discussion

5.1 Is N_2 Fixation Relevant in the Gulf of Aqaba?

In this study we implemented and optimized a series of models with different assumptions about N_2 fixation in the Gulf of Aqaba. The models range from one neglecting N_2 fixation to another assuming that, in addition to two autotrophic diazotroph groups, N_2 fixation can occur in the entire water column (i.e., independent of light availability). While the models are very similar in their abilities to replicate chlorophyll and PO_4, model H3 performed the best in reproducing the observed pattern of deep-NO_3 accumulation and O_2. Overall, all models that consider N_2 fixation accumulate nitrogen at different rates, as they enrich the nitrogen content of detritus, which is then remineralized at depth over time.

The best model performance was obtained with two groups of autotrophic organisms and a group of heterotrophic organisms (H3). A model without explicit N_2 fixation, but in the absence of sediment denitrification, also increases the accumulation of deep NO_3 in a similar fashion as version H2 but not sufficiently high to match the observations. This suggests that N_2 fixation in the area must exceed denitrification rates. In the models with denitrification, the average sediment denitrification flux is 0.25 ± 0.46 mmol N m^{-2} d^{-1}, with a maximum value of 3.01 mmol N m^{-2} d^{-1}. These values at are the lower end of a global compilation of sediment denitrification rates by Fennel et al. (2009), which have a mean of 2.2 mmol N m^{-2} d^{-1} and maximum values exceeding 10 mmol N m^{-2} d^{-1}.

Observations from the Gulf of Aqaba exhibit an excess of nitrogen that contrasts to exterior waters from the Arabian Sea and Indian Ocean, which are considered net nitrogen sink regions (Gruber and Sarmiento, 1997). There are too few reported values of dissolved inorganic nitrogen-to-phosphorus ratios for the Red Sea region from Bab-el-Mandeb to the Strait of Tiran to provide a complete idea of its spatial distribution; however, the limited available information supports our conclusions. Naqvi et al. (1986) found a significant difference in N:P ratios between surface incoming and sub-surface outflowing waters at Bab-el-Mandeb, concluding that N_2 fixation was a process required to explain these anomalies in the nitrogen budget. Higher nitrogen concentrations have been observed in the Red Sea in comparison to the Arabian Sea and Indian Ocean, where a strong deficit of nitrogen develops as losses due to denitrification exceed the input of newly fixed nitrogen (Burkill et al., 1993; Naqvi, 1994; Gruber and Sarmiento, 1997; Morrison et al., 1998, 1999). Close to the entrance of the Persian Gulf, average excess phosphate has been estimated to be around 5 mmol m^{-3} at all depths and seasons reported, with maximum excess values on the order of 8 mmol m^{-3} (Gruber and Sarmiento, 1997). Thus, it has been hypothesized that limited deep-water exchange at Bab-el-Mandeb allows waters of the Red Sea outside of the Gulf of Aqaba to acquire different characteristics from inflowing Arabian Sea waters (Naqvi et al., 1986). Our model results support this hypothesis and suggest that N_2 fixation is key for the formation of the distinct biochemical characteristics in the Gulf of Aqaba. Considering the regional context, our models suggest that, despite low rates, N_2 fixation is necessary to explain the nitrogen vertical distribution in the Gulf of Aqaba, and the interannual accumulation of deep nitrate during years with weak convection.

5.2 How does N_2 fixation Contribute to Primary Production?

In this section we discuss the contribution of N_2 fixation to primary production in the Gulf of Aqaba, and our quantitative estimates of N_2 fixation with respect to global rates (Figures 8-9). Our estimates of surface primary productivity agree with those reported by Iluz et al. (2009) for March-April of 2008. However, our models overestimate surface primary productivity values in 2010 when compared to those reported by Rahav et al. (2015). On average, our best-performing model version yields an annual primary production rate of 304±56.9 g C m^{-2} yr^{-1} (H3). This rates is higher than previously published annual averages, which range from 80 g C m^{-2} y^{-1} (Levanon-Spanier et al., 1979) to 170 g C m^{-2} y^{-1} (Iluz, 1991), whereas more recent unpublished primary production estimates at IUI range between 141 and 197 g C m^{-2} y^{-1} (pers. comm. Y. Shaked).

The ratio of new to total primary production (f-ratio) in our experiments ranges from 15% to 80%. Maximum f-ratios are estimated in January and February due to significant contributions from deep NO_3, whereas f-ratios are at their minimum during stratified conditions (June – August). Our best-performing model version, H3, estimates a summer minimum f-ratio 0.22. The average f-ratio for all scenarios is 0.47. This agrees with published estimates for the Gulf of Aqaba of 0.5 during the stratified period as determined from a nitrate-diffusion model (Badran et al., 2005).

Total annual N_2 fixation rates from our best-performing model versions (H3 and H3a) are similar to high estimates reported for other regions (Capone and Carpenter, 1982; Michaels et al., 1996; Lee et al., 2002), while those obtained in the other model experiments are within the range of values reported for the Gulf of Aqaba. The intensity of winter mixing has a minor effect on N_2 fixation rates; the largest effect occurred in H3a where N_2 fixation increased by 15% after deep winter mixing. Based on our best-performing model version (H3), we estimate that 10% to 14% of the total primary production is supported by N_2 fixation.

5.3 Are Heterotrophic N_2 Fixers Important?

In contrast to previous models (e.g., Hood et al., 2001; Fennel et al., 2002; Monteiro et al., 2010; Moore et al., 2004), our model version H3 relaxes the assumption of light dependence for diazotrophy, through the inclusion of heterotrophic diazotrophs in addition to two groups of autotrophic diazotrophs. This model improves the representation of NO_3 and O_2 at depth (Figures 3, 6). Changes in deep NO_3 can be explained through the enrichment of detritus, while changes in O_2 may reflect the additional sink of O_2 at depth due to the heterotrophic group. All model versions with heterotrophic organisms also match observed estimates of N_2 fixation in deep waters of the Gulf of Aqaba, and without them N_2 fixation rates below the DCM are underestimated (Figure 8). Heterotrophic N_2 fixation also impacts total N_2 fixation (Figure 9).

There is growing evidence of non-cyanobacterial N_2 fixation in aphotic waters (Benavides et al., 2017; Moisander et al., 2017). For instance, N_2 fixation rates in mesopelagic and abyssopelagic waters down to 2000 m (Fernandez et al., 2011; Bonnet et al., 2013; Loescher et al., 2014) have been attributed to non-cyanobacterial organisms, including proteobacteria (Turk-Kubo et al., 2014). These *nif*H-expressing heterotrophic phylotypes can be as abundant as unicellular cyanobacterial groups and dominate the deep and dark zones of the water column (Church et al. 2005; Langlois et al., 2005; Riemann et al., 2010). Genetic evidence and rate estimates from the Gulf of Aqaba suggest that *nif*H expressing heterotrophic proteobacteria α and γ may explain the correlation of bacterial productivity rates with N_2 fixation rates (Rahav et al., 2013; 2015). Aside from the Gulf

of Aqaba, aphotic N_2 fixation and *nif*H gene expression have also been reported in the Baltic Sea (Farnelid et al., 2013), Arabian Sea (Jayakumar et al., 2012) and Mediterranean Sea (Rahav et al., 2013).

Despite the importance of heterotrophic diazotrophs in our model, the simulated colonial diazotroph blooms are responsible for the highest N_2 fixation rates, so they are a necessary model aspect to achieve resemblance with the observations. This is in line with evidence of extensive blooms of *Trichodesmium spp.* being responsible for the high N_2 fixation rates observed in the Arabian Sea and Red Sea (Capone et al., 1998; Post et al., 2002; Foster et al., 2009). In the northern Gulf of Aqaba, colonies and free trichomes of *Trichodesmium spp.* are found throughout the year down to 100 m depth (Post et al., 2002). Ephemeral blooms of *T. erythraeum* and *T. thiebautii* have been documented near the coast of Eilat (Post et al., 2002, Gordon et al., 1994; Kimor and Golandsky, 1977). However, massive blooms are rare in the Gulf of Aqaba (Foster et al., 2009; Mackey et al., 2007, pers. Communication Berman-Frank) and the model probably overestimates the contribution of *Trichodesmium spp.*'s annual blooming to total N_2 fixation rates, as seen in the much larger surface N_2 fixation rates generated by H2 and H3. As new observational information is collected, further model refinements may be necessary to better reflect the actual contribution of different diazotrophic groups in the Gulf of Aqaba.

5.4 Limitations and Uncertainties

The one-dimensional nature of our physical setting, which neglects the contribution of horizontal advection to the vertical structure of simulated tracers, can be considered a limitation of this study. This simplification is, however, necessary to perform model calibration and test multiple model structures at a manageable computational expense. We applied temperature and salinity nudging to ensure accurate representation of the vertical density structure. Comparison of the simulated vertical structure with and without nudging shows that this correction has negligible effects on deep waters, where the effects of N_2 fixation are the most relevant. This is consistent with the existing literature about circulation of the Gulf of Aqaba, which describes how geomorphology and bathymetry limit water exchange between the Gulf of Aqaba and the Red Sea to the upper 300 m (Wolf-Vecht et al., 1992; Biton and Gildor, 2011). It is, therefore, unlikely that horizontal transport could explain the observed accumulation of deep NO_3. Nevertheless, transport of nitrogen-enriched sub-surface waters from the Gulf of Aqaba towards the exterior may dampen the long-term accumulation of nitrogen (Figure 7).

There are other sources of nitrogen that were not explored in the present study that we discuss here briefly. For instance, we did not include contributions to N_2 fixation by diatom-diazotroph

associations, which are significant in other regions. While diatom-diazotroph associations have been detected in the Gulf of Aqaba, they are not as abundant as unicellular diazotrophs, *Trichodesmium* and proteobacteria (Kimor et al. 1992, Foster et al., 2009, pers. Comm Berman-Frank). In general, due to the oligotrophic characteristics of the region, small phytoplankton species (<8 μm) contribute more than 90% of the chlorophyll-a standing stock (Lindell & Post 1995, Yahel et al. 1998). Dinoflagellates and diatoms together correspond to less than 5% of the phytoplankton biomass, except during ephemeral diatom blooms during spring when they can account for nearly 50% of the total biomass (Al-Najjar et al., 2006).

Another source of nitrogen which could significantly affect this region is atmospheric deposition, as the Gulf receives considerable dust input from the surrounding deserts. Recently, it has been shown that atmospheric dust input does not correlate with chlorophyll variability in surface waters of the Gulf of Aqaba (Torfstein and Kienast, 2018). A previous study suggested that atmospheric deposition of nitrogen could support over 10% of surface primary production in the region, based on measurements of local aerosol composition and a dust deposition model (Chen et al., 2007). However, this estimate had a relatively large uncertainty due to errors associated with the deposition flux calculation and the temporal variability in dust flux (Chen et al., 2007). Moreover, very low nitrogen concentrations and N:P ratios lower than Redfield from the surface down to 80 m were observed during the same time period (Foster et al., 2009). Therefore, the role of atmospheric nitrogen inputs remains uncertain.

In contrast to the improvement in NO_3 distributions with the addition of heterotrophic diazotrophs, all other model versions exhibit similar underestimation of deep total PO_4. This suggests that their structure lacks a process affecting this nutrient. Several processes can independently affect PO_4. Prior knowledge suggests that seasonal advection of nitrogen depleted surface waters may contribute to changes in PO_4, while having minimum impact on NO_3 concentrations. Other relevant missing processes include variable plankton stoichiometry and PO_4 remineralization. Our sensitivity analysis highlighted that varying the planktonic N:P ratios has an evident effect on PO_4 but does not affect NO_3 distributions significantly. A related effect was noticed when comparing fixed constant vs. variable plankton nutrient uptake stoichiometry in a model of the central Baltic Sea: seasonal particulate organic C:N ratios were very similar, but C:P diverged (Kreuss et al., 2015). Moreover, preferential PO_4 remineralization may more directly increase deep-PO_4 concentrations as organic matter is decomposed in the bottom layers. The inclusion of this process in a model of the North Atlantic Ocean improved the representations of biogeochemical characteristics of the area and increased the N_2 fixation rates obtained by the

model (Monteiro and Follows, 2012). Variable nitrogen allocation within diazotrophs also improved the performance of a model representing annual cycles in the North Atlantic sub-tropical gyre (Fernández-Castro et al., 2016). Further investigation of this process is recommended.

Finally, an intrinsic caveat that should accompany all biological models is the uncertainty associated with parameter values (Denman, 2003). We reduced this uncertainty by using systematic parameter optimization for a subset of parameters, yet others were kept fixed at a priori assumed values. In the case of N:P ratios, a simple sensitivity analysis suggests that they do not affect conclusions with respect to inorganic dissolved nitrogen but modify phosphorus concentrations. In the case of parameters related to diazotrophs, these are largely unconstrained by the observations we used, which are typically available in long-term data sets (i.e., Chl-a, NO_3, PO_4, O_2). We emphasize that results in this study are experimental, testing the effects of assumptions about diazotroph behaviour rather than modifying such behaviour by subjectively tuning parameter values.

6 Conclusions

We implemented and optimized biogeochemical models that represent a range of different assumptions about diazotrophy in a 700 m-deep pelagic station from the northern Gulf of Aqaba. Our model results demonstrate the importance of N_2 fixation in replicating the observed water-column-integrated nitrogen and oxygen inventories. The model without N_2 fixation is unable to replicate the observed vertical structure of inorganic nitrogen. The models that include diazotrophs significantly modify this variable by increasing the fraction of remineralized nitrogen from organic matter decomposition. The effect of N_2 fixation on O_2 distributions depends on the type of nitrogen fixer. N_2 fixation by autotrophic-photosynthetic organisms increases oxygen concentrations, while heterotrophic organisms decrease deep-oxygen due to increased respiration and organic matter decomposition. The observed vertical structure of NO_3 and oxygen is reproduced best with a model that includes heterotrophic, and colonial and unicellular autotropic diazotrophs suggesting that heterotrophic N_2 fixation is necessary to explain the observed excess nitrogen at this location. N_2 fixation assumptions do not affect PO_4 concentrations significantly, but they are affected by assumptions about N:P ratios of organic matter. The N_2 fixation rates simulated by this model are similar to the highest observational estimates from the Gulf of Aqaba. Aphotic N_2 fixation is simulated to occur at lower rates than maximum autotrophic N_2 fixation yet occurs continuously over a large portion of the water column. This suggests that heterotrophic

diazotrophs set a background rate of N_2 fixation in the ocean that should be considered further in global estimates and biogeochemical models.

Acknowledgements We thank the Inter University Institute for Marine Sciences in Eilat and A. Genin for providing the observational data, and E. Landou, B. Lazar (BL), S. Kienast, H. Gildor, E. Boss, A. Torfstein, and M. Lewis (ML) for insightful discussions. We gratefully acknowledge financial support from the Schulich Marine Studies Initiative to KF, IBF, ML, and BL.

References

Amon, R.M.W. and Benner, R.: Photochemical and microbial consumption of dissolved organic carbon and dissolved oxygen in the Amazon River system, Geochimica et Cosmochimica Acta, 60, 10, 1783-1792, 1996.

Badran, M.I., Rasheed, M., Manasrah, R. and Al-Najjar, T.: Nutrient flux fuels the summer primary productivity in the oligotrophic waters of the Gulf of Aqaba, Red Sea, Oceanologia 47, 1, 2005.

Ben-Sasson, M., Brenner, S. and Paldor, N.: Estimating air–sea heat fluxes in semienclosed basins: The case of the Gulf of Elat (Aqaba), J. Phys. Oceanogr., 39, 185–202, 2009.

Berman, T., Paldor, N. and Brenner, S.: The seasonality of tidal circulation in the Gulf of Elat, Isr. J. Earth Sci., 52, 2003.

Berman, T., Paldor, N. and Brenner, S.: Simulation of wind-driven circulation in the Gulf of Elat (Aqaba), J. Mar. Syst., 26, 349–365. doi:10.1016/S0924-7963(00)00045-2, 2000.

Berman-Frank, I., Cullen, J.T., Shaked, Y., Sherrell, R.M. and Falkowski, P.G.: Iron availability, cellular iron quotas, and nitrogen fixation in Trichodesmium, Limnol. Oceanogr., 46, 1249–1260, 2001.

Biton, E., Silverman, J. and Gildor, H.: Observations and modeling of a pulsating density current, Geophys. Res. Lett., 35, 14, 2008.

Biton, E. and Gildor, H.: The general circulation of the Gulf of Aqaba (Gulf of Eilat) revisited: The interplay between the exchange flow through the Straits of Tiran and surface fluxes, J. Geophys. Res., 116, C8, doi:10.1029/2010JC006860, 2011.

Bonnet, S., Bethelot, H., Turk-Kubo, K., Fawcett, S., Rahav, E., L'Helguen, S. and Berman-Frank, I.: Dynamics of N_2 fixation and fate of diazotroph-derived nitrogen in a low-nutrient, low-

chlorophyll ecosystem: results from the VAHINE mesocosm experiment (New Caledonia), Biogeosciences, 13, 2653–2673. doi:10.5194/bg-13-2653-2016, 2016.

Bonnet, S., Dekaezemacker, J., Turk-Kubo, K.A., Moutin, T., Hamersley, R.M., Grosso, O., Zehr, J.P. and Capone, D.G.: Aphotic N_2 fixation in the eastern tropical South Pacific Ocean, PloS one, 8, 12, e81265, 2013.

Burchard, H., Bolding, K. and Villarreal, M.: GOTM- a general ocean turbulence model. Theory, applications and test cases (Technical Report No. EUR 18745 EN), European Commission, 1999.

Burkill, P.H., Mantoura, R.F.C. and Owens, N.J.P.: Biogeochemical cycling in the northwestern Indian Ocean: a brief overview, Deep Sea Res. Part II Top. Stud. Oceanogr., 40, 643–649, 1993.

Capone, D.G., Burns, J.A., Montoya, J.P., Subramaniam, A., Mahaffey, C., Gunderson, T., Michaels, A.F. and Carpenter, E.J.: Nitrogen fixation by *Trichodesmium spp.*: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean, Glob. Biogeochem. Cycles, 19, 2, 2005.

Capone, D.G. and Carpenter, E.J.: Nitrogen fixation in the marine environment, Science, 217, 1140–1142, 1982.

Capone, D.G., Subramaniam, A., Montoya, J.P., Voss, M., Humborg, C., Johansen, A.M., Siefert, R.L. and Carpenter, E.J.: An extensive bloom of the N_2-fixing cyanobacterium *Trichodesmium erythraeum* in the central Arabian Sea, Mar. Ecol. Prog. Ser., 172, 281–292, 1998.

Capone, D.G., Zehr, J.P., Paerl, H.W., Bergman, B. and Carpenter, E.J.: Trichodesmium, a globally significant marine cyanobacterium, Science, 276, 1221–1229, 1997.

Carlson, D.F., Fredj, E. and Gildor, H.: The annual cycle of vertical mixing and restratification in the Northern Gulf of Eilat/Aqaba (Red Sea) based on high temporal and vertical resolution observations, Deep Sea Res. Part Oceanogr. Res. Pap., 84, 1–17, 2014.

Carlson, D.F., Fredj, E., Gildor, H., Biton, E., Steinbuck, J.V., Monismith, S.G. and Genin, A.: Observations of tidal currents in the northern Gulf of Eilat/Aqaba (Red Sea), J. Mar. Syst., 102, 14–28, 2012.

Carpenter, E.J. and McCarthy, J.J.: Nitrogen fixation and uptake of combined nitrogenous nutrients by Oscillatoria (Trichodesmium) thiebutii in the western Sargasso Sea, Limnol. Oceanogr., 20, 389–401, 1975.

Carpenter, E.J., Montoya, J.P., Burns, J.A., Mulholland, M.R. and Subramaniam, A., Capone, D.G.: Extensive bloom of a N_2-fixing diatom/cyanobacterial association in the tropical Atlantic Ocean, Mar. Ecol. Prog. Ser., 185, 273–283, 1999.

Carpenter, E.J. and Roenneberg, T.: The marine planktonic cyanobacteria *Trichodesmium spp.*: photosynthetic rate measurements in the SW Atlantic Ocean, Mar. Ecol. Prog. Ser., 267–273, 1995.

Chase, Z., Paytan, A., Johnson, K.S., Street, J. and Chen, Y.: Input and cycling of iron in the Gulf of Aqaba, Red Sea, Glob. Biogeochem. Cycles, 20, 3, 2006.

Chen, Y., Mills, S., Street, J., Golan, D., Post, A., Jacobson, M. and Paytan, A.: Estimates of atmospheric dry deposition and associated input of nutrients to Gulf of Aqaba seawater, J. Geophys. Res. Atmosphere, 112, D4, 2007.

Church, M.K., Jenkinks, Bethany D., Karl, D.M. and Zehr, J.: Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the oligotrophic North Pacific Ocean. Aquat. Microb. Ecol. 38, 3–14, 2005.

Denman, K.L.: Modelling planktonic ecosystems: parameterizing complexity, Prog. Oceanogr., 57, 429–452, 2003.

Deutsch, C., Sigman, D.M., Gruber, N. and Dunne, J.P.: Spatial coupling of nitrogen inputs and losses in the ocean, Nature, 445, 163–168. doi:10.1038/nature05392, 2007.

Dutkiewicz, S., Ward, B.A., Monteiro, F.M. and Follows, M.: Interconnection nitrogen fixers and iron in the Pacific Ocean: Theory and numerical simulations, Glob. Biogeochem. Cycles, 26, 1, 2012.

Dutkiewicz, S., Morris, J., Follows, M.J., Scott, J., Levitan, O., Dyhrman, S.T. and Berman-Frank, I.: Impact of ocean acidification on the structure of future phytoplankton communities, Nature Climate Change, 5, 11, 1002-1006, doi: 10.1038/NCLIMATE2722, 2015.

Enríquez, S.C.M.D., Duarte, C.M. and Sand-Jensen, K.A.J.: Patterns in decomposition rates among photosynthetic organisms: the importance of detritus C: N: P content, Oecologia, 94, 4, 457-471, 1993.

Eppley, R.W. and Peterson, B.J.: Particulate organic matter flux and planktonic new production in the deep ocean, Nature, 282, 677, 1979.

Falkowski, P.G.: Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean, Nature, 387, 272–274, 1997.

Fahnenstiel, G.L., McCormick, M.J., Lang, G.A., Redalje, D.G., Lohrenz, S.E., Markowitz, M., Wagoner, B. and Carrick, H.J.: Taxon-specific growth and loss rates for dominant phytoplankton populations from the northern Gulf of Mexico, Mar. Ecol. Prog. Ser., Oldendorf ,117, 1, 229-239, 1995.

Farnelid, H., Bentzon-Tilia, M., Andersson, A.F., Bertilsson, S., Jost, G., Labrenz, M., Jürgens, K. and Riemann, L.: Active nitrogen-fixing heterotrophic bacteria at and below the chemocline of the central Baltic Sea, The ISME Journal, 7, 1413–1423, doi:10.1038/ismej.2013.26, 2013.

Fennel, K., Brady, D., DiToro, D., Fulwiler, R.W., Gardner, W.S., Giblin, A., McCarthy, M.J., Rao, A., Seitzinger, S., Thouvenot-Korppoo, M. and Tobias, C.: Modeling denitrification in aquatic sediments, Biogeochemistry, 93, 159–178, 2009.

Fennel, K., Follows, M. and Falkowski, P.G.: The co-evolution of the nitrogen, carbon and oxygen cycles in the Proterozoic, Ocean. *Am. J. Sci.*, 305, 526-545, 2005.

Fennel, K., Hu, J., Laurent, A., Marta-Almeida, M. and Hetland, R.: Sensitivity of hypoxia predictions for the northern Gulf of Mexico to sediment oxygen consumption and model nesting, J. Geophys. Res. Oceans, 118, 990–1002, 2013.

Fennel, K., Spitz, Y.H., Letelier, R. and Abbott, M.R.: A deterministic model for N_2 fixation at stn. ALOHA in the subtropical North Pacific Ocean, Deep-Sea Res. II, 49, 149–174, 2002.

Fennel, K., Wilkin, J., Levin, J., Moisan, J., O'Reilly, J.E. and Haidvogel, D.: Nitrogen cycling in the Middle Atlantic Bight: Results from a three-dimensional model and implications for the North Atlantic nitrogen budget, Glob. Biogeochem. Cycles 20, 14, doi:10.1029/2005GB002456, 2006.

Fennel**,** K., Losch, M., Schröter, J., and Wenzel M.: Testing a marine ecosystem model: Sensitivity analysis and parameter optimization**,** *J. Mar. Syst.,* 28, 1, 45-63, 2001.

Fernandez, C., Farías, L., & Ulloa, O. (2011). Nitrogen fixation in denitrified marine waters. PloS one, 6(6), e20539.

Fernández-Castro, B., Pahlow, M., Mouriño-Carballido, B., Marañón, E., & Oschlies, A. (2016). Optimality-based Trichodesmium diazotrophy in the North Atlantic subtropical gyre. Journal of Plankton Research, 38(4), 946-963.

Foster, R.A., Paytan, A. and Zehr, J.P.: Seasonality of N2 fixation and nifH gene diversity in the Gulf of Aqaba (Red Sea), Limnol. Oceanogr., 54, 219–233, 2009.

Fuller, N.J., West, N.J., Marie, D., Yallop, M., Rivlin, T., Post, A.F. and Scanlan, D.J.: Dynamics of community structure and phosphate status of picocyanobacterial populations in the Gulf of Aqaba, Red Sea, Limnol. Oceanogr., 50, 363–375, 2005.

Geider, R.J., MacIntyre, H.L. and Kana, T.M.: Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a: carbon ratio to light, nutrient-limitation and temperature, Mar. Ecol. Prog. Ser., 148, 187–200, 1997.

Gifford, D.J., Fessenden, L.M., Garrahan, P.R. and Martin, E.: Grazing by microzooplankton and mesozooplankton in the high-latitude North Atlantic Ocean: Spring versus summer dynamics, J. of Geophys. Res. Oceans, 100, C4, 6665-6675, 1995.

Genin, A. and Paldor, N.: Changes in the circulation and current spectrum near the tip of the narrow, seasonally mixed Gulf of Elat, Isr. J. Earth. Sci., 47, 87–92, 1998.

Gordon, N., Angel, D.L., Neori, A., Kress, N. and Kimor, B.: Heterotrophic dinoflagellates with symbiotic cyanobacteria and nitrogen limitation in the Gulf of Aqaba, Mar. Ecol. Prog. Ser., 83–88, 1994.

Gregg, W.W.: Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model, J. Mar. Syst., 69, 205–225, 2008.

Gruber, N. and Galloway, J.N.: An Earth-system perspective of the global nitrogen cycle, Nature, 451, 293–296, 2008.

Gruber, N. and Sarmiento, J.L.: Global patterns of marine nitrogen fixation and denitrification, Glob. Biogeochem. Cycles, 11, 235–266, 1997.

Haug, G.H., Pedersen, T.F., Sigman, D.M., Calvert, Nielsen, B. and Peterson, L.C.: Glacial/interglacial variations in production in nitrogen fixation in the Cariaco Basin during the last 580 kyr, Paleoceanography, 13, 427–432, 1998.

Hong, H., Shen, R., Zhang, F., Wen, Z., Chang, S., Lin, W., ... & Shi, D. (2017). The complex effects of ocean acidification on the prominent N2-fixing cyanobacterium Trichodesmium. Science, 356(6337), 527-531.

Hood, R., Bates, N.R., Capone, D.G. and Olson, D.B.: Modeling the effect of nitrogen fixation on carbon and nitrogen fluxes BATS, Deep-Sea Res. II, 48, 1609–1648, 2001.

Houck, C.R., Joines, J.A. and Kay, M.G.: A genetic algorithm for function optimization: A Matlab implementation (Technical Report No. NCSU-IE-TR-95-09), North Carolina State University, Raleigh, NC, 1995.

Hutchins, D.A., Fu, F.-X., Zhang, Y., Warner, M.E., Feng, Y., Portune, K., Bernhardt, P. and Mullholland, M.R.: CO_2 control of Trichodesmium N_2 fixation, photosynthesis, growth rates, and elemental ratios: Implications for past, present, and future ocean biogeochemistry, Limnol. Oceanogr., 52, 1293–1304, 2007.

Iluz, D.: Primary production of phytoplankton in the northern Gulf of Eilat, Red Sea, MSc thesis, Bar-Ilan University, Ramat-Gan (in Hebrew with English abstract), 1991.

Iluz, D., Dishon, G., Capuzzo, E., Meeder, E., Astoreca, R., Montecino, V., Znachor, P., Ediger, D. and Marra, J.: Short-term variability in primary productivity during a wind-driven diatom bloom in the Gulf of Eilat (Aqaba), Aquat. Microb. Ecol., 56, 205–215, 2009.

Karl, D.: Nutrient dynamics in the deep blue sea, TRENDS Microbiol., 10, 410–418, 2002.

Jayakumar, A., Al-Rshaidat, M.M.D., Ward, B.B. and Mulholland, M.R.: Diversity, distribution, and expression of diazotroph *nifH* genes in oxygen-deficient waters of the Arabian Sea *FEMS Microbiology Ecology*, 82, 3, 597–606,https://doi.org/10.1111/j.1574-6941.2012.01430.x, 2012.

Kerbrat, A.S., Darius, H.T., Pauillac, S., Chinain, M., and Laurent, D.: Detection of ciguatoxin-like and paralysing toxins in *Trichodesmium* spp. from New Caledonia lagoon, Mar. Pollut. Bull., 2010, 61, 360–366, 2010.

Kerbrat, A.S., Amzil, Z., Pawlowiez, R., Golubic, S., Sibat, M., Darius, H.T, Chinain, M. and Laurent, D.: First Evidence of Palytoxin and 42-Hydroxy-palytoxin in the Marine Cyanobacterium *Trichodesmium.*, Mar. Drugs, 9, 543-560, doi:10.3390/md9040543, 2011.

Kimor, B. and Golandsky, B.: Microplankton of the Gulf of Elat: aspects of seasonal and bathymetric distribution, Mar. Biol., 42, 55–67, 1977.

Klinker, J., Reiss, Z., Kropach, C., Levanon, I., Harpaz, H., Halicz, E. and Assaf, G.: Observations on the circulation pattern in the Gulf of Elat (Aqaba), Red Sea, Isr. J. Earth Sci., 25, 85–103, 1976.

Kranz, S. A., Levitan, O., Richter, K. U., Prášil, O., Berman-Frank, I. and Rost, B.: Combined effects of CO2 and light on the N2-fixing cyanobacterium Trichodesmium IMS101: physiological responses, Plant Physiology, 154, 1, 334-345, 2010.

Kreus, M., Schartau, M., Engel, A., Nausch, M., & Voss, M. (2015). Variations in the elemental ratio of organic matter in the central Baltic Sea: Part I—Linking primary production to remineralization. Continental Shelf Research, 100, 25-45.

Kuhn, A.M., Fennel, K. and Mattern, J.P.: Model investigations of the North Atlantic spring bloom initiation, Prog. Oceanogr., 138, 176–193, 2015.

Labiosa, R.G., Arrigo, K.R., Genin, A., Monismith, S.G. and van Dijken, G.: The interplay between upwelling and deep convective mixing in determining the seasonal phytoplankton dynamics in the Gulf of Aqaba: Evidence from SeaWiFS and MODIS, Limnol. Oceanogr., 48, 2355–2368, 2003.

Landry, M.R., Hasset, R.P., Fagerness, V., Downs, J. and Lorenzen, C.J.: Effect of food acclimation on assimilation efficiency of *Calanus pacificus*, Limnol. and Oceanogra., 29, 361–364, 1984.

Lazar, B., Erez, J., Silverman, J., Rivlin, T., Rivlin, A., Dray, M., Meeder, E. and Iluz, D.: Recent environmental changes in the chemical-biological oceanography of the Gulf of Aqaba (Eilat), Aqaba-Eilat Improbable Gulf Environment, Biodiversity and Preservation, Magnes Press, Jerusalem, 49–62, 2008.

Lee, K., Karl, D.M., Wanninkhof, R. and Zhang, J., Z.: Global estimates of net carbon production in the nitrate-depleted tropical and subtropical oceans, Geophys. Res. Lett., 29, 19, 2002.

Levanon-Spanier, I., Padan, E. and Reiss, Z.: Primary production in a desert-enclosed sea-the Gulf of Elat (Aqaba), Red Sea, Deep Sea Res. Part Oceanogr. Res. Pap., 26, 673–685, 1979.

Lima, I.D. and Doney, S.C.: A three-dimensional, multinutrient, and size-structured ecosystem model for the North Atlantic, Glob. Biogeochem. Cycles, 18, 3, 2004.

Luo, Y.W., Doney, S.C., Anderson, L.A., Benavides, M., Berman-Frank, I., Bode, A., Bonnet, S., Boström, K.H., Böttjer, D., Capone, D.G. and Carpenter, E.J.: Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates, Earth System Science Data, 4, 1, 47-73, 2012.

Mackey, K. R., Labiosa, R. G., Calhoun, M., Street, J. H., Post, A. F., and Paytan, A.: Phosphorus availability, phytoplankton community dynamics, and taxon-specific phosphorus status in the Gulf of Aqaba, Red Sea, Limnol. and Oceanogra., 52, 2, 873-885, 2007.

Mackey, K.R., Rivlin, T., Grossman, A.R., Post, A.F. and Paytan, A.: Picophytoplankton responses to changing nutrient and light regimes during a bloom, Mar. Biol., 156, 1531–1546, 2009.

Manasrah, R., Raheed, M. and Badran, M.I.: Relationships between water temperature, nutrients and dissolved oxygen in the northern Gulf of Aqaba, Red Sea, Oceanologia, 48, 2, 2006.

Meeder, E., Mackey, K.R., Paytan, A., Shaked, Y., Iluz, D., Stambler, N., Rivlin, T., Post, A.F. and Lazar, B.: Nitrite dynamics in the open ocean-clues from seasonal and diurnal variations, Mar. Ecol. Prog. Ser., 453, 2012.

Michaels, A.F., Olson, D., Sarmiento, J.L., Ammerman, J.W., Fanning, K., Jahnke, R., Knap, A.H., Lipschultz, F. and Prospero, J.M.: Inputs, losses and transformations of nitrogen and phosphorus in the pelagic North Atlantic Ocean, in: Nitrogen Cycling in the North Atlantic Ocean and Its Watersheds, Springer, 181–226, 1996.

Moisander, P., Beinart, R.A., Hewson, I., White, A.E., Johnson, K.S., Carlson, C.A., Montoya, J.P. and Zehr, J.P.: Unicellular cyanobacterial distribution broaden the oceanic N_2 fixation domain, Science, 327, 1512–1514, 2010.

Moisander, P., Benavides, M., Bonnet, S. and Berman-Frank, I.: Chasing after Non-cyanobacterial Nitrogen Fixation in Marine Pelagic Environments, Front. Microbiol., 8, 1736, doi: https://doi.org/10.3389/fmicb.2017.01736, 2017.

Monismith, S.G. and Genin, A.: Tides and sea level in the Gulf of Aqaba (Eilat), J. Geophys. Res. Oceans, 109, 2004.

Monteiro, F.M., Follows, M. and Dutkiewicz, S.: Distribution of diverse nitrogen fixers in the global ocean, Glob. Biogeochem. Cycles, 24, 3, 2010.

Monteiro, F. M., & Follows, M. J.: On nitrogen fixation and preferential remineralization of phosphorus. Geophysical Research Letters, 39(6), 2012.

Montoya, J.P., Holl, C.M., Zehr, J.P., Hansen, A., Villareal, T.A. and Capone, D.G.: High rates of N_2 fixation by unicellular diazotrophs in the oligotrophic Pacific Ocean, Nature, 420, 1027–1032, 2004.

Moore, J.K., Doney, S. and Lindsay, K.: Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model, Glob. Biogeochem. Cycles, 18, 4, 2004.

Moore, J.K., Doney, S.C., Kleypas, J.C., Glover, D.M. and Fung, I.Y.: An intermediate complexity marine ecosystem model for the global domain, Deep Sea Res., Part II, 49, 403–462, 2002.

Morrison, J.M., Codispoti, L.A., Gaurin, S., Jones, B., Manghnani, V. and Zheng, Z.: Seasonal variation of hydrographic and nutrient fields during the US JGOFS Arabian Sea Process Study, Deep Sea Res. Part II Top. Stud. Oceanogr., 45, 2053–2101, 1998.

Morrison, J.M., Codispoti, L.A., Smith, S.L., Wishner, K., Flagg, C., Gardner, W.D., Gaurin, S., Naqvi, S.W.A., Manghnani, V., Prosperie, L. and Gundersen, J.S.: The oxygen minimum zone in the Arabian Sea during 1995, Deep Sea Res. Part II Top. Stud. Oceanogr., 46, 1903–1931, 1999.

Naqvi, S.W.A.: Denitrification processes in the Arabian Sea, Proc. Indian Acad. Sci.-Earth Planet. Sci., 103, 279–300, 1994.

Naqvi, S. W. A., Hansen, H.P. and Kureishy, T.W.: Nutrient-uptake and regeneration ratios in the Red-Sea with reference to the nutrient budgets, Oceanologica acta, 9, 3, 271-275, 1986.

Nielsen, S.L.: Size-dependent growth rates in eukaryotic and prokaryotic algae exemplified by green algae and cyanobacteria: comparisons between unicells and colonial growth forms, J. Plankton Res. 28, 5, 489–498, 2006.

O'Neil, J.M. and Roman, M.R.: Ingestion of the cyanobacterium *Trichodesmium spp.* by pelagic harpacticoid copepods Macrosetella, Miracia and Oculosetella, Ecology and Morphology of Copepods, Springer, 235–240, 1994.

Pomeroy, L.R. and Wiebe, W.J.: Temperature and substrates as interactive limiting factors for marine heterotrophic bacteria, Aquat. Microb. Ecol., 23, 187–204, 2001.

Post, A. F., Dedej, Z., Gottlieb, R., Li, H., Thomas, D.N., El-Absawi, M., El-Naggar, A., El-Gharabawi, M. and Sommer, U.: Spatial and temporal distribution of *Trichodesmium spp.* in the stratified Gulf of Aqaba, Red Sea, Mar. Ecol. Prog. Ser., 239, 241–250, 2002.

Rahav, E., Bar-Zeev, E., Ohayion, S., Elifantz, H., Belkin, N., Herut, B., Mulholland, M.R. and Berman-Frank, I.R.: Dinitrogen fixation in aphotic oxygenated marine environments, Front. Microbiol., 4, 227, 2013.

Rahav, E., Herut, B., Mulholland, M.R., Belkin, N., Elifantz, H. and Berman-Frank, I.: Heterotrophic and autotrophic contribution to dinitrogen fixation in the Gulf of Aqaba, Mar. Ecol. Prog. Ser., 522, 67–77, 2015.

Riemann, L., Farnelid, H. and Steward, G.F.: Nitrogenase genes in non-cyanobacterial plankton: prevalence, diversity and regulation in marine waters, Aquat. Microb. Ecol., 61, 235-247, 2010.

Rubin, M., Berman-Frank, I. and Shaked, Y.: Dust-and mineral-iron utilization by the marine dinitrogen-fixer Trichodesmium, Nature Geoscience, 4, 8, 529-534, 2011.

Schartau, M. and Oschlies, A.: Simultaneous data-based optimization of a 1D-ecosystem model at three locations in the North Atlantic: Part I - Method and parameter estimates, J. Mar. Res., 61, 765–793, 2003.

Silverman, J. and Gildor, H.: The residence time of an active versus a passive tracer in the Gulf of Aqaba: A box model approach, J. Mar. Syst., 71, 159–170, 2008.

Smayda, T. and Bienfang, P.K., Suspension properties of various phyletic groups of phytoplankton and tintinnids in an oligotrophic, subtropical system, Mar. Ecol., 4, 289–300, 1983.

Smith, E. L. (1936). Photosynthesis in relation to light and carbon dioxide. Proceedings of the National Academy of Sciences of the United States of America, 22(8), 504.

Tande, K.S. and Slagstad, D.: Assimilation efficiency in herbivorous aquatic organisms - The potential of the ratio method using 14C and biogenic silica as markers, Limnol. and Oceanogra., 30, 1093–1099, 1985.

Veldhuis, M.J.W, Timmermans, K.R., Croot, P. and van der Wagt B.: Picophytoplankton; a comparative study of their biochemical composition and photosynthetic properties, J. of Sea Res., 53, 1, 7-24, 2005.

Ward, B.A., Dutkiewicz, S., Moore, C.M. and Follows, M.: Iron, phosphorus, and nitrogen supply ratios define the biogeography of nitrogen fixation, Limnol. Oceanogr., 58, 2059–2075, 2013.

Ward, B.A., Friedrichs, M.A.M., Anderson, T.R. and Oschlies, A.: Parameter optimization techniques and the problem of underdetermination in marine biogeochemical models, J. Mar. Syst., 81, 34–43, 2010.

Wolf-Vecht, A., Paldor, N. and Brenner, S.: Hydrographic indications of advection/convection effects in the Gulf of Elat, Deep Sea Res. Part Oceanogr. Res. Pap., 39, 1393–1401, 1992.

Zarubin, M., Lindemann, Y. and Genin, A.: The dispersion-confinement mechanism: Phytoplankton dynamics and the spring bloom in a deeply-mixing subtropical sea, Prog. in Oceanogr., 155, 13-27, 2017.

Zehr, J.P.: Nitrogen fixation by marine cyanobacteria, Curr. Trends Microbiol., 19, 4, 162–173, 2011.

Zehr, J.P., Bench, S.R., Carter, B.J., Hewson, I., Niazi, F., Shi, T., Tripp, H.J. and Affourtit, J.: Globally distributed uncultivated oceanic N_2-fixing cyanobacteria lack ocygenic photosystem II, Science, 322, 1110–1112, 2008.

Zehr, J.P., Carpenter, E.J. and Villareal, T.A.: New perspectives on nitrogen-fixing microorganisms in tropical and subtropical oceans, Curr. Trends Microbiol., 8, 2, 68-73, 2000.

Zehr, J.P., Mellon, M.T. and Zani, S.: New Nitrogen-Fixing Microorganisms Detected in Oligotrophic Oceans by Amplification of Nitrogenase (nifH) Genes. Appl. Environ. Microbiol. 64, 3444–3450, 1998.

Zehr, J.P. and Ward, B.B.: Nitrogen cycling in the ocean: new perspectives on processes and paradigms. Appl. Environ. Microbiol. 68, 1015–1024, 2002.

Zehr, J.P., Waterbury, J.B., Turner, P., Montoya, J.P., Omoregie, E., Steward, G.F., Hansen, A. and Karl, D.M.: Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. Nature 412, 635–637, 2001.

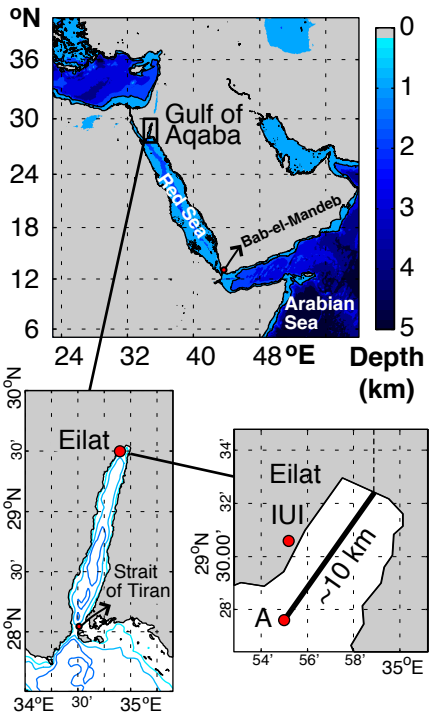

Figure 1: Map of study area showing the location of monitoring stations and geographic references.

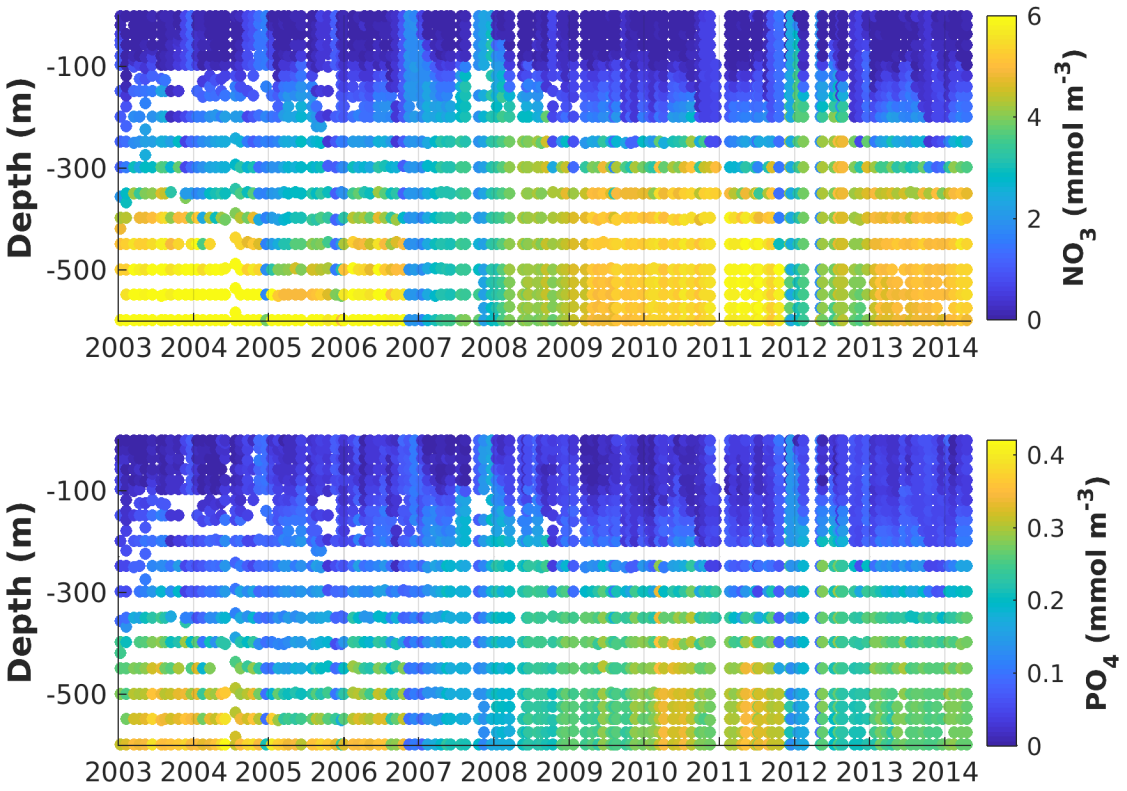

Figure 2: Observed vertical nitrate and phosphate distributions at Station A (Gulf of Aqaba) from 2004 to 2014. Tick marks are placed on the April 1st of every year. (modified colorscale)

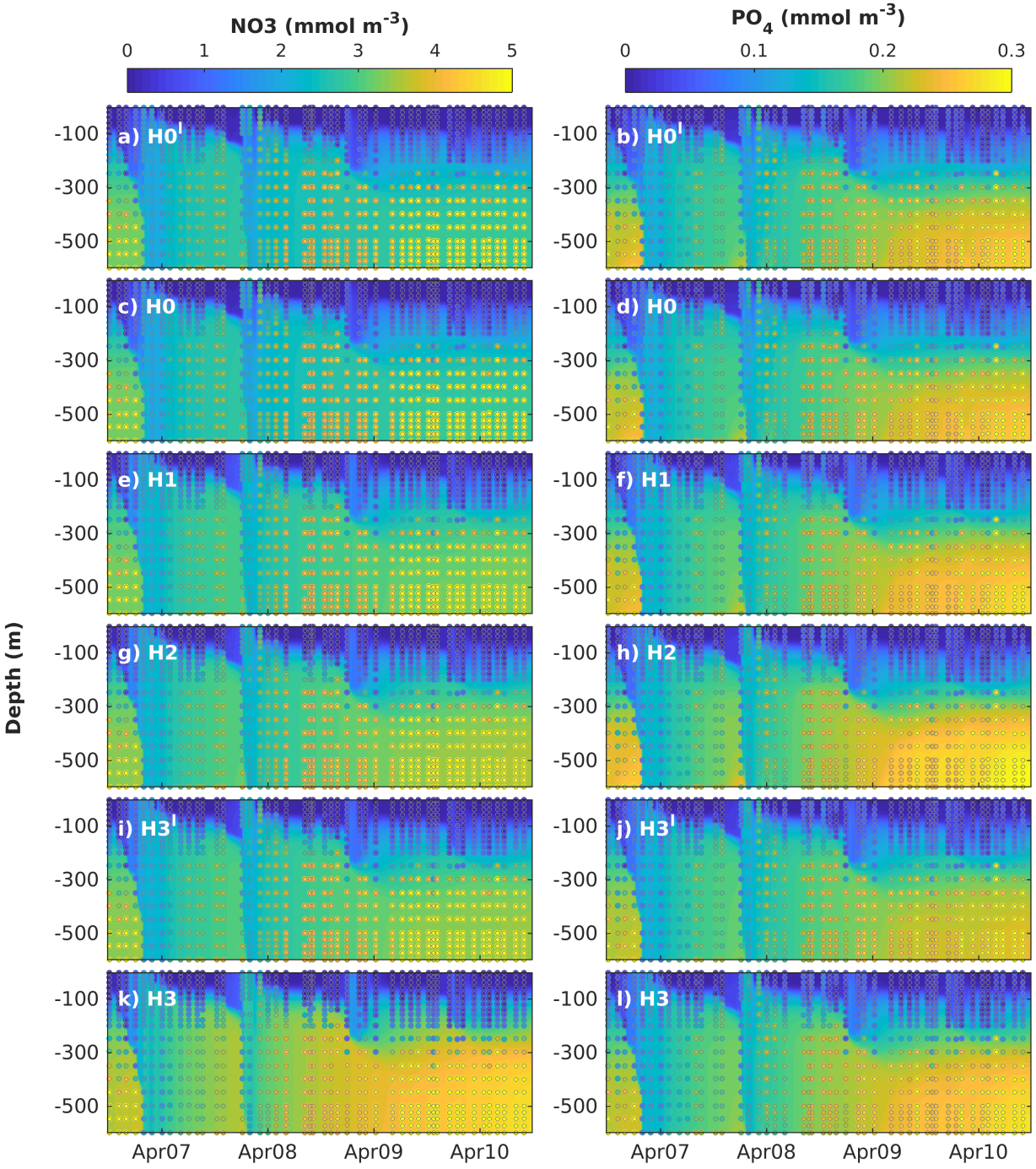

Figure 3: Observed (coloured circles) and simulated (background) NO3 and PO4 using model versions H0 (no nitrogen fixers), H1 (generic autotrophic fixer), H2 (unicellular and colonial autotrophic fixers), H3' (non-fixing heterotrophic, and unicellular and colonial autotrophic fixers), H3 (heterotrophic, and unicellular and colonial autotrophic fixers). The spin-up period is not shown. (modified colorscale)

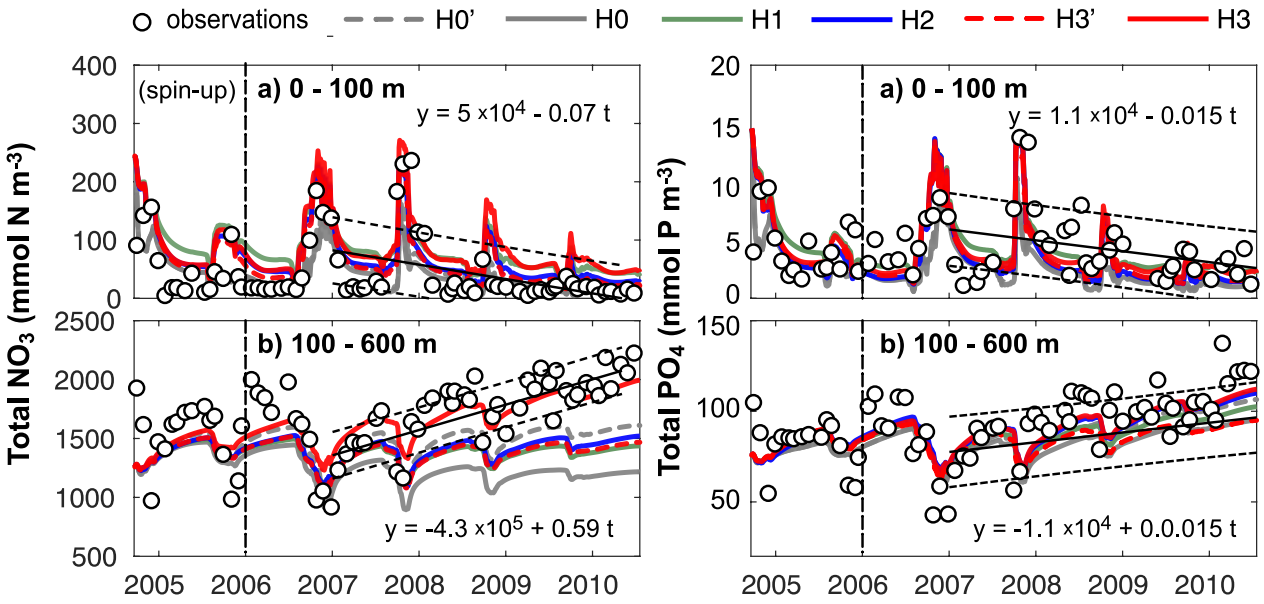

Figure 4: Observed and simulated vertically integrated NO3 and PO4 between 0 – 100 m and between 100 – 600 m using model versions H0 (no nitrogen fixers), H0' (no sediment denitrification – no fixers), H1 (generic autotrophic fixer), H2 (unicellular and colonial autotrophic fixers), H3' (non-fixing heterotrophic, and unicellular and colonial autotrophic fixers), H3 (heterotrophic, and unicellular and colonial autotrophic fixers). Tick marks are placed on the April 1st of every year.

-

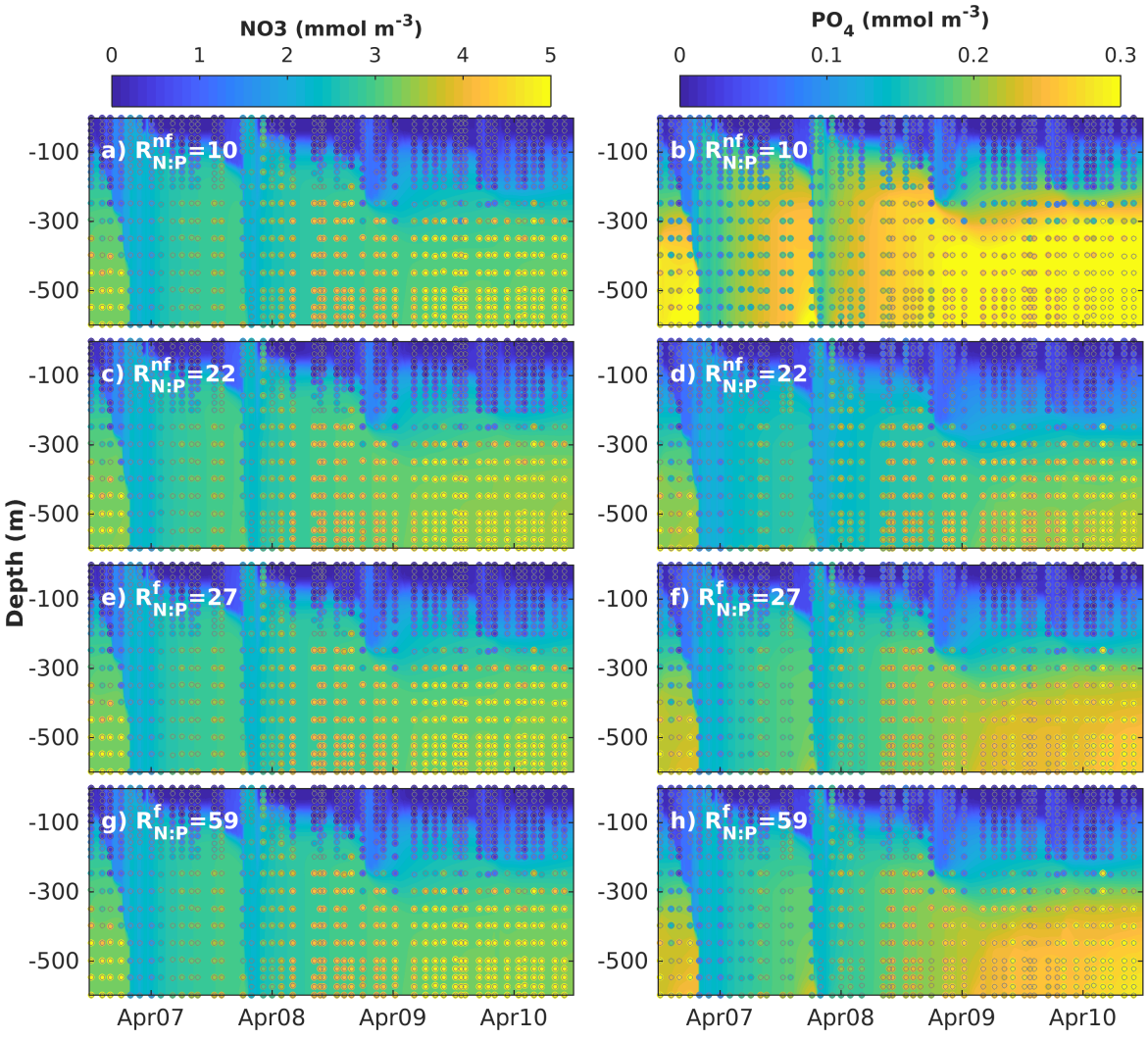

Figure 5: Sensitivity experiments modifying the fixed N:P ratios in non-fixing organisms ($R^{nf}_{N:P}$) and N_2 fixing organisms ($R^{f}_{N:P}$). Observed (coloured circles) and simulated (background) NO_3 and PO_4 using model versions H1 (generic autotrophic fixer). The spin-up period is not shown.

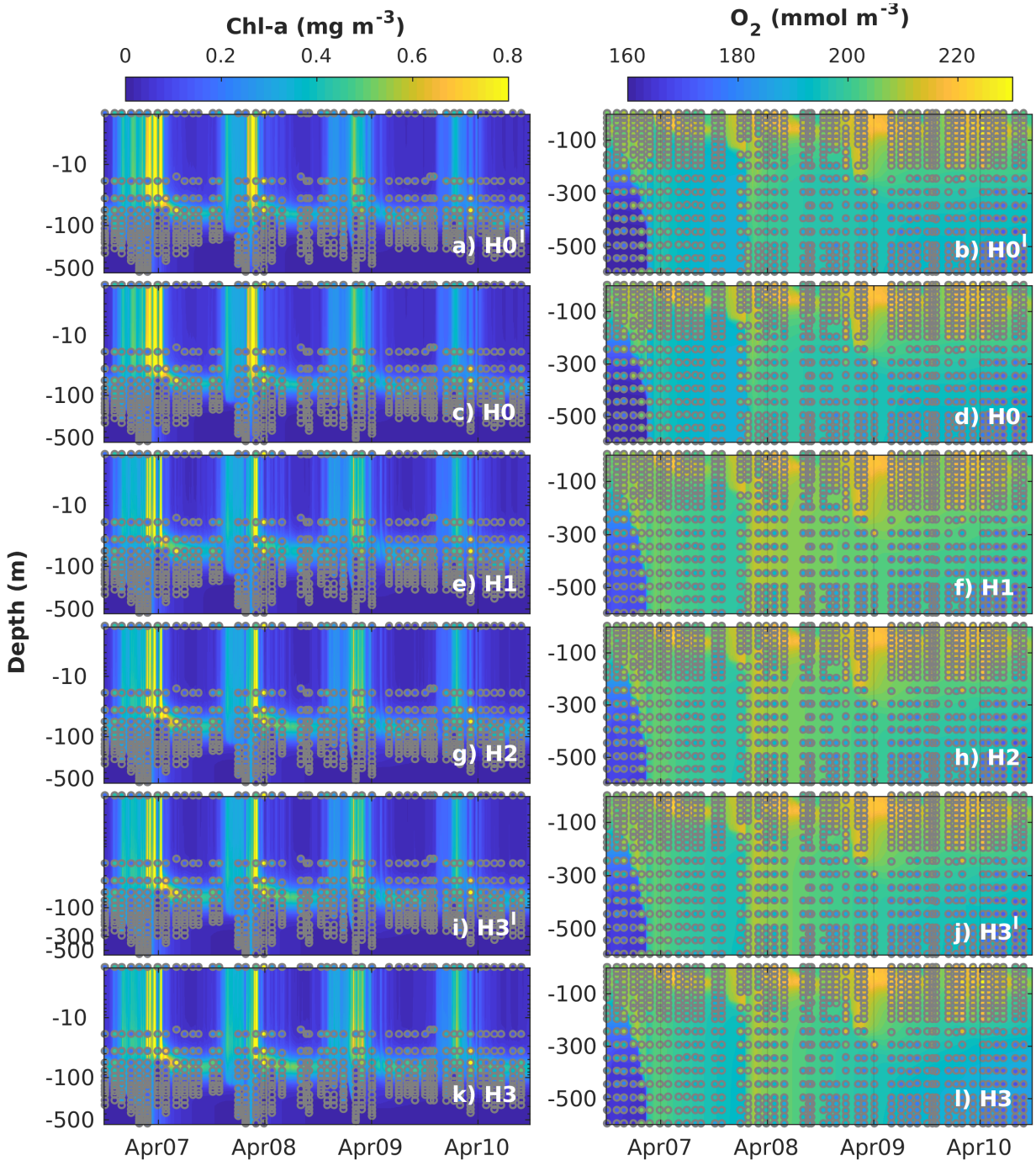

Figure 6: Observed (coloured circles) and simulated (background) Chl-a and O_2 using model versions H0' (no sediment denitrification – no fixers), H0 (no nitrogen fixers), H1 (generic autotrophic fixer), H2 (unicellular and colonial autotrophic fixers), H3' (unicellular and colonial autotrophic fixers and heterotrophic non-fixer), H3 (non-fixing heterotrophic, and unicellular and colonial autotrophic fixers), and H3 (heterotrophic, and unicellular and

colonial autotrophic fixers). Vertical scale in the Chl-a subplots is logarithmic to exaggerate the surface. The spin-up period is not shown. (modified colorscale)

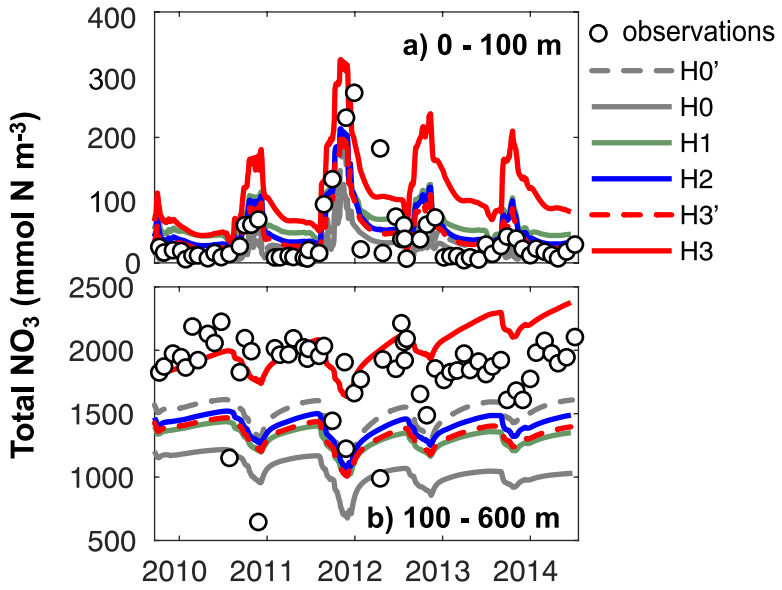

Figure 7: Observed (circles) and simulated (lines) total nitrate in the surface and deep waters at Station A during the model validation period from 2010 to 2014. Tick marks are placed on the April 1st of every year.

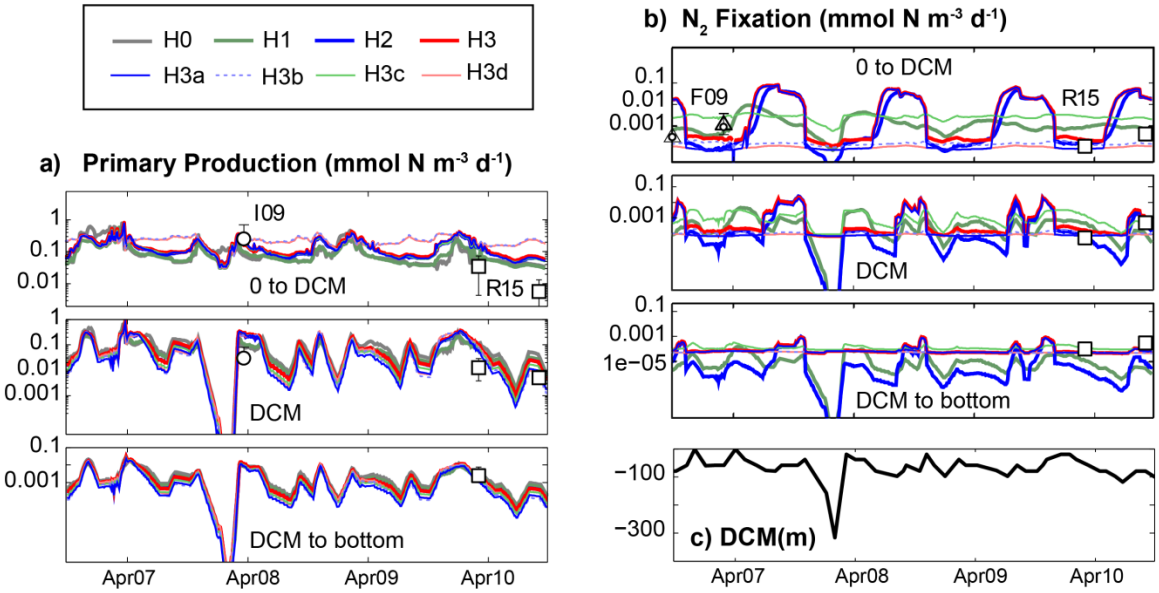

Figure 8: Comparison of previously reported in situ measurements and model results of primary production (a) and N_2 fixation rates (b), averaged at three depth levels. Depth levels

are from the surface to the Deep Chlorophyll Maximum, at the DCM and below it. (c) DCM estimated from observed Chl-a profiles at Station A. I09, F09 and R15 refer to Iluz et al., (2009), Foster et al., (2009), and Rahav et al., (2015), respectively. Control versions H0' and H3' are not shown.

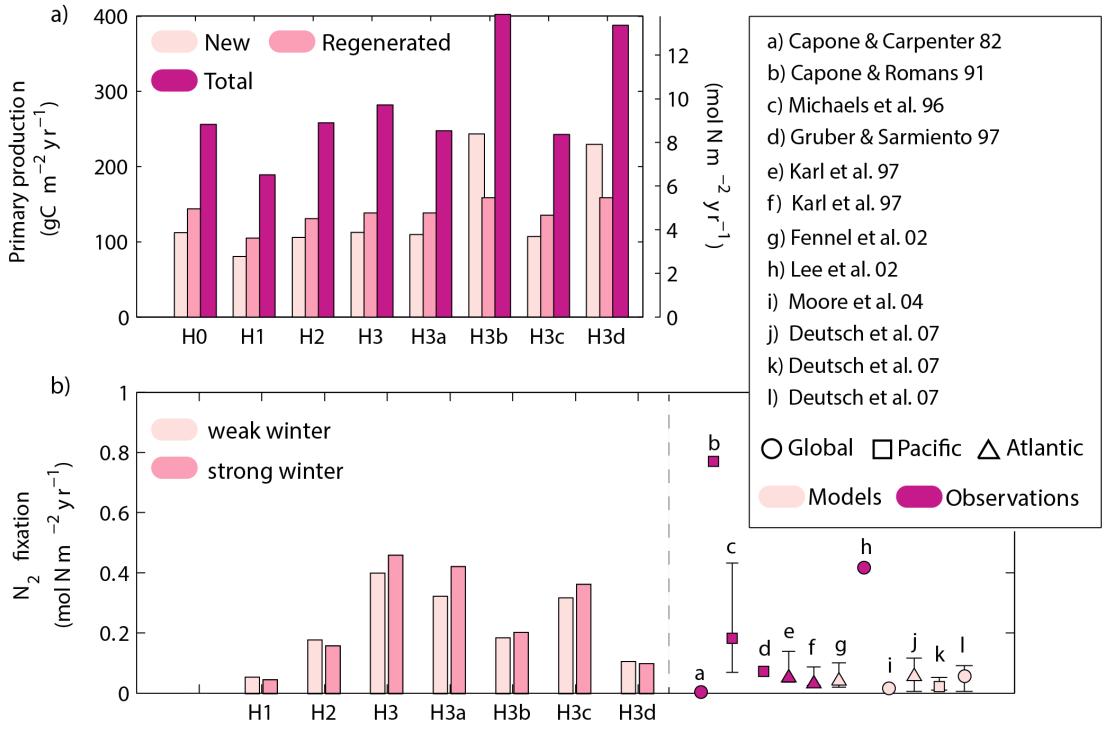

Figure 9: Simulated new, regenerated, and total primary production (a) and N_2 fixation rates (b) obtained by the different simulations (see Table 1for key to different simulations). A summary of previous estimates of N_2 fixation rates in observational and model studies is included in (b). Redfield C:N ratio is used to transform the model results nitrogen units to carbon units. Control versions H0' and H3' are not shown.

Table 1: Summary of model version characteristics and assumptions about diazotroph groups. N = no diazotrophs; A = autotrophic diazotrophs; H = heterotrophic diazotrophs. Checkmarks (✓) represent presence of a model characteristic / functional diazotrophic group in the model; dashes represent their absence.

Model Version	Characteristics / Diazotrophs Groups	Denitrification	Generic (A)	Unicellular (A)	Colonial (A)	Heterotrophic (H)
H0'	N	−	-	-	-	-
H0	N	✓	-	-	-	-
H1	A	✓	✓	-	-	-
H2	A	✓	-	✓	✓	-
H3'	A H	✓		✓	✓	**Non-fixing**
H3	A H	✓	-	✓	✓	✓
H3a	A H	✓	-	-	✓	✓
H3b	A H	✓	-	✓	-	✓
H3c	A H	✓	✓	-	-	✓
H3d	H	✓	-	-	-	✓
	Diazotrophs characteristics:					
	Inorganic phosphorus uptake		✓	✓	✓	✓
	Organic phosphorus uptake		-	-	-	✓
	Light growth limitation		✓	✓	✓	-
	Temperature dependent maximum growth rate		✓	✓	✓	✓
	Minimum temperature limit for growth (20^{o}C)		-	-	✓	-
	Predation		-	✓	-	-

Table 2: Parameters used in the base biogeochemical model (H0), including minimum and maximum parameters ranges based on the literature. Parameters values followed by * were obtained by the optimization.

Parameters	Value	Range	Description	Units	References
μ_{Phy}^{0}	0.76*	0.1 – 3	Reference phytoplankton maximum growth rate at T = 0ºC	d^{-1}	a, b, c
$k_{Phy}^{NO_3}$	0.05	0.01 – 0.5	Phytoplankton NO_3 uptake half-saturation	mmol m^{-3}	d, e
$k_{Phy}^{NH_4}$	0.1*	0.01 – 0.5	Phytoplankton NH_4 uptake half-saturation	mmol m^{-3}	d, e
k_{Phy}^{DIP}	0.004*	0.001 – 0.5	Phytoplankton DIP uptake half-saturation	mmol m^{-3}	a, f, g
α_{Phy}	0.1*	0.01 – 0.125	Phytoplankton, initial slope of photosynthetic response	molC $gChl^{-1}$ $(W\ m^{-2})^{-1}\ d^{-1}$	d, h
m_{Phy}	0.1	0.01 – 0.2	Phytoplankton mortality rate	d^{-1}	d
g_{Phy}^{max}	1.16*	0.1 - 4	Zooplankton maximum grazing rate	d^{-1}	b, i
k_{Zoo}^{Phy}	0.5*	0.01 – 0.5	Square zooplankton grazing half-saturation	$(mmol\ m^{-3})^2$	d, e
l_{BM}	0.011*	0.01 – 0.15	Zooplankton base metabolic rate	d^{-1}	d
l_E	0.1	0.05 – 0.35	Zooplankton excretion rate	d^{-1}	d
m_Z	0.35*	0.02 - 0.35	Zooplankton mortality rate	d^{-1}	d

τ	0.1	0.01 - 25	Small detritus aggregation rate	d^{-1}	d, e
θ_{Phy}^{max}	0.142*	0.015 – 0.15	Maximum chlorophyll to carbon ratio	mg Chl (mg C)$^{-1}$	h
β	0.74*	0.25 – 0.75	Zooplankton assimilation efficiency	non-dim.	j, k
r_{DOM}	0.2	0.05 – 0.5	DOM remineralization rate	d^{-1}	l
r_D	0.01	0.005 – 0.15	Detritus remineralization rate	d^{-1}	l, m
n_{max}	0.3*	0.01 – 0.35	Nitrification rate	d^{-1}	d, e
k_I	0.1	0.01 – 0.5	Half-saturation radiation for nitrification inhibition	Wm^{-2}	d
I_{th}	0.0095	0.005 – 0.01	Radiation threshold for nitrification inhibition	Wm^{-2}	d
w_{Phy}	0.1	0.01 – 1	Vertical sinking velocity for non-fixing phytoplankton	md^{-1}	n
w_{DL}	-4.44*	0.01 – 25	Vertical sinking velocity for large detritus	md^{-1}	d

a. Fennel et al. (2002) b. Fahnenstiel et al. (1995) c. Veldhuis et al. (2005) d. Fennel et al. (2006) e. (Lima and Doney (2004) f. Ward et al. (2013) g. Moore, et al.(2002) h. Geider et al. (1997) i. Gifford et al. (1995) j. Landry et al. (1984) k. Tande and Slagstad (1985) l. Amon and Benner (1996) m. Enríquez et al. (1993) n. Smayda and Bienfang (1983)

Table 3: Diazotrophs parameters and re-calibrated non-fixing phytoplankton parameters for each model version. H0 = no N2 fixers; H1 = generic autotrophic diazotrophs; H2 = unicellular and colonial cyanobacteria; H3 = heterotrophs, unicellular and colonial cyanobacteria. Superscripts in the parameter descriptions denote the corresponding literature references used to define the generic diazotroph parameter values within a realistic range. In order to tease apart the effect of different diazotrophs niches or behaviour, unicellular, colonial and heterotrophic diazotrophs maintain the same parameter values as the generic organism. An exception is the slightly lower reference growth rate used for heterotrophic diazotrophs.

Model version:	H0	H1	H2	H3	Units	Description
μ_{Phy}^{0}	0.76	2.20	1.5	1.5	d^{-1}	Reference phytoplankton maximum growth rate at T = 0ºC
θ_{Phy}^{max}	0.022	0.076	0.076	0.05	mg Chl (mg C)$^{-1}$	Maximum chlorophyll to carbon ratio – non fixing phytoplankton
$k_{Phy}^{NH_4}$	0.076	0.076	0.076	0.076	mmol m^{-3}	Phytoplankton NH_4 uptake half-saturation
k_{Phy}^{DIP}	0.001	0.015	0.015	0.015	mmol m^{-3}	Phytoplankton DIP uptake half-saturation
m_{Phy}	0.1	0.06	0.06	0.06	d^{-1}	Phytoplankton mortality rate
g_{Phy}^{max}	1.16	4.0	1.95	1.95	d^{-1}	Zooplankton maximum grazing rate
β	0.36	0.7	0.7	0.7	non-dim.	Zooplankton assimilation efficiency
$\mu_{G_F}^{0}$	-	0.25	-	-	d^{-1}	Reference generic diazotrophs maximum growth rate at T = 0ºC[a-j]

Model version:	H0	H1	H2	H3	Units	Description
$k_{G_F}^{DIP}$	-	0.001	-	-	mmol m^{-3}	Generic diazotrophs DIP uptake half-saturation[d]
θ_F^{max}	-	0.053	-	-	mg Chl (mg C)$^{-1}$	Maximum chlorophyll to carbon ratio – generic diazotrophs[l]
α_{G_F}	-	0.01	-	-	molC gChl^{-1} (W m^{-2})$^{-1}$ d^{-1}	Generic diazotrophs, initial slope of photosynthetic response[a, d, g, k]
m_{G_F}	-	0.18	-	-	d^{-1}	Generic diazotrophs mortality rate[d]
l_{G_F}	-	0.05	-	-	d^{-1}	Generic diazotrophs respiration rate[d]
$\mu_{U_F}^0$	-	-	0.25	0.25	d^{-1}	Reference unicellular cyanobacteria maximum growth rate at T = 0ºC
$k_{U_F}^{DIP}$	-	-	0.004	0.004	mmol m^{-3}	Unicellular cyanobacteria DIP uptake half-saturation
$\theta_{U_F}^{max}$	-	-	0.053	0.053	mg Chl (mg C)$^{-1}$	Maximum chlorophyll to carbon ratio – unicellular cyanobacteria
α_{U_F}	-	-	0.05	0.05	molC gChl^{-1} (W m^{-2})$^{-1}$ d^{-1}	Unicellular cyanobacteria, initial slope of photosynthetic response
m_{U_F}	-	-	0.20	0.2	d^{-1}	Unicellular cyanobacteria mortality rate
l_{U_F}	-	-	0.05	0.05	d^{-1}	Unicellular cyanobacteria respiration rate
$g_{U_F}^{max}$	-	-	0.2	0.2	d^{-1}	Zooplankton maximum grazing rate on unicellular cyanobacteria
$k_{Zoo}^{U_F}$	-	-	0.001	0.001	(mmol m^{-3})^2	Square zooplankton grazing half-saturation on unicellular cyanobacteria

Model version:	H0	H1	H2	H3	Units	Description
$\mu_{C_F}^0$	-	-	0.25	0.25	d^{-1}	Reference colonial cyanobacteria maximum growth rate at T = 0ºC
$k_{C_F}^{DIP}$	-	-	0.004	0.004	mmol m^{-3}	Colonial cyanobacteria DIP uptake half-saturation
$\theta_{C_F}^{max}$	-	-	0.053	0.053	mg Chl (mg C)$^{-1}$	Maximum chlorophyll to carbon ratio – colonial cyanobacteria
α_{C_F}	-	-	0.05	0.05	molC $gChl^{-1}$ (W m^{-2})$^{-1}$ d^{-1}	Colonial cyanobacteria, initial slope of photosynthetic response
m_{C_F}	-	-	0.18	0.05	d^{-1}	Colonial cyanobacteria mortality rate
l_{C_F}	-	-	0.18	0.05	d^{-1}	Colonial cyanobacteria respiration rate
$\mu_{H_F}^0$	-	-	-	0.2	d^{-1}	Reference heterotrophs maximum growth rate at T = 0ºC[m]
$k_{H_F}^{DIP}$	-	-	-	0.001	mmol m^{-3}	Heterotrophs DIP uptake half-saturation
$k_{H_F}^{DS}$	-	-	-	0.001	mmol m^{-3}	Heterotrophs organic phosphorus uptake half-saturation
m_{H_F}	-	-	-	0.2	d^{-1}	Heterotrophs mortality rate
l_{H_F}	-	-	-	0.05	d^{-1}	Heterotrophs respiration rate

a. Moore, et al., (2004); b. Ward et al., (2013); c. Hood et al., (2001); d. Fennel et al., (2002); e. Capone et al., (1997); f. Berman-Frank et al., (2001); g. Hutchins et al., (2007); h. Kranz et al., (2010); i. Hong et al., (2017); k. Geider et al., (1997); l. Fennel et al., (2006); m. Pomeroy and Wiebe, (2001)

Table 4 Root-mean-square-errors between observations and corresponding simulated variables. Observations between 2005 and 2010 were used during model calibrations (i.e., assimilated). Observations between 2011 and 2014 are used for independent model

validations (non-assimilated). Control versions H0' and H3' are not shown.

0-100 m								
	2005 – 2010 (assimilated)				**2011 – 2014 (non-assimilated)**			
	NO_3	**PO_4**	**CHL**	**O_2**	**NO_3**	**PO_4**	**CHL**	**O_2**
H0	0.71	0.04	0.15	7.57	0.60	0.04	0.16	6.39
H1	0.77	0.04	0.14	6.99	0.66	0.04	0.15	7.08
H2	0.78	0.04	0.14	6.96	0.75	0.04	0.14	6.66
H3	1.04	0.05	0.14	7.35	1.50	0.05	0.13	6.22
H3a	1.04	0.06	0.12	7.10	1.41	0.09	0.14	6.47
H3b	1.91	0.05	0.14	7.94	2.15	0.05	0.16	8.13
H3c	1.01	0.06	0.12	7.05	1.06	0.08	0.14	6.55
H3d	1.60	0.05	0.19	7.91	1.78	0.05	0.19	8.21
100 – 600 m								
	NO_3	**PO_4**	**CHL**	**O_2**	**NO_3**	**PO_4**	**CHL**	**O_2**
H0	1.53	0.05	0.08	15.54	2.26	0.06	0.06	17.34
H1	1.43	0.05	0.07	15.09	2.02	0.05	0.06	21.70
H2	1.29	0.05	0.07	14.17	1.56	0.04	0.06	17.57
H3	1.05	0.05	0.07	13.28	0.89	0.05	0.06	10.03
H3a	1.12	0.05	0.07	13.42	0.93	0.07	0.06	11.98
H3b	1.26	0.10	0.07	18.29	1.51	0.14	0.06	29.99
H3c	1.14	0.05	0.07	14.37	1.18	0.05	0.06	16.16
H3d	1.41	0.11	0.14	18.39	1.85	0.14	0.13	30.38