Response to reviewer

We appreciate the time this reviewer has taken to comment on our manuscript for the second time. We have considered their suggestions and answer them in detail below.

[reviewer comments in blue]

The authors have done a reasonable job in responding to my previous review. However, I do have a couple of concerns related to two of the more substantive changes which the authors implemented in response to comments previously raised by both myself and the other reviewers.

Major comments

Firstly, many of the reviewers, myself included, commented on the poor agreement between the modelled and the observed N* distributions in the previous version. In response to this the authors have removed direct comparison between observed and modelled N* and instead rely on comparisons between model results and observations of N and P separately. However, the use of N* (or effectively N:P ratios), which will reflect the decoupling between the cycling of N and P (see e.g. Sarmiento and Gruber 1997 and numerous subsequent usage), would still seem like the most obvious and direct constraint on the N cycle in general and N2 fixation in particular within the model. As such I personally found the decision to remove this comparison, rather than, for example, to directly address the prior weakness by investigating how the added sensitivity analyses (see below) influenced model reproduction of N* (and/or N:P ratios) to be a somewhat unsatisfactory response.

Response: The decision to remove N* figures and discussion was taken after the explicit suggestion of one of our first-round reviewers. Also, in response to those comments, we had added significant discussion about limitations and additional sources of N and P not included in the model, which could improve model performance. In this new version, we have reintroduced the observed and simulated N* figures and accompanying discussion. However, we feel that the addition of further figures analyzing the sensitivity of N* to planktonic N:P ratios does not add significant information—to some extent, this information can be inferred from the effects on N and P. As N:P ratios mainly affect P concentrations, we would expect a deviation of N* toward excess phosphate when phosphate is increased (Fig 7 b, h in the revised manuscript) and an enhancement of excess nitrate when phosphate decreases (Fig 7 d in the revised manuscript).

Additionally, I was pleased to see the authors perform a sensitivity analysis investigating how changes in the assumed ratios of N:P within both non-N2 fixing phytoplankton and nitrogen fixers influenced their results. However, this sensitivity analysis appeared quite limited. In particular, I was unclear why this analysis was only performed for the H1 version of the model? Also, I note that the sensitivity analysis did not really cover the parameter space of different assumed N:P ratios for phytoplankton and nitrogen fixers very thoroughly, as they were varied independently around their initial assumed values. Wouldn't it be more appropriate to perform

a sensitivity analysis over a matrix of values fully investigating the parameter space?

As an example of how the two points above might interact, I would particularly have liked to see whether a sensitivity analysis of the assumed N:P ratios within the H3 version of the model might influence the ability of this most complex structure to simultaneously reproduce P and N observations, as indicated by N* values (or maybe simply N:P ratios in the subsurface).

Response: We acknowledge that investigation of the effects of planktonic stoichiometry is relevant for a better understanding of N* variability. Nevertheless, an extensive exploration of the model sensitivity to N:P is beyond of the intended scope of the present manuscript and may distract the narrative from our main arguments. One-at-a-time variations in parameter values are a frequently used exploratory tool and we have used sensible value ranges for both fixing and non-fixing phytoplankton types (10-28 and 19 – 59, respectively). In the manuscript, we only show a few examples of the results (both increasing and decreasing the initial N:P value), but 6 different values within those ranges were tested generating similar results. While N:P ratios may vary temporarily outside the value ranges we tested, using those values as fixed constant ratios would generate unrealistic responses.

More sophisticated methods to investigate sensitivity are often the subject of independent analysis. They may involve techniques to optimally sample the parameter space and/or the use of statistical model emulators to minimize computational time required by such extensive evaluations of the parameter space. We estimate that a minimum of ~625 model runs would be required for the sensitivity analysis suggested by the reviewer.

Additional more minor comments

Page 4, Line 20: this statement could perhaps be more specific given the different perturbations/versions of each of the models (H0-H3) now investigated.

Response: We modified the text from *"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"* to *"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 main alternative ecosystem model versions and six model subversions with minor variations**"*

Page 5, Line 27: 'diazotrophs' rather than 'diazotrophics'
Response: Thank you, corrected.

Page 7, Lines 5-14: I'm still not clear why this optimisation method couldn't be performed for the N:P ratios for the phytoplankton and diazotrophs? (see also Page 7, Line 28)
Response: We applied the optimization method to parameters that often exhibit large value ranges, as an alternative to calibrate the model subjectively. While we acknowledge that N:P

ratios can be variable, there is more consensus about N:P ratios for N2 fixing and non-fixing phytoplankton when used as fixed values in models. Here we kept these values to maintain consistency with common assumptions in previous models. The method could be applied to N:P ratios; however, in order to obtain more reliable solutions to these parameters, PON and POP data would be needed.

Page 9, Line 25 onwards. Related to second major comment above, in the absence of any further analysis, some rationale for why this sensitivity analysis was only performed for H1 and only around the initially assumed values (rather than more fully exploring the parameter space), should be provided here.

Response: As all diazotrophic groups share some characteristics of the phosphate uptake parameterizations, the behaviour of the generic diazotroph is indicative of potential effects in the most extensive model versions in a simplified context. We have added this comment on the main text.