Anonymous Referee #1

General Comments

R: A premise of this paper is that two version of a model with different assumptions about diazotrophy and which have similar patterns of nitrogen fixation in the current day, have very different responses in a future climate scenario. This is a useful comment to make in terms of modeling of climate change impacts. The fact that different assumptions can lead to similar patterns of diazotrophy has indeed been seen before (e.g. Landolfi et al 2015). However, that future changes lead to different outcomes has not been documented to my knowledge. Given this premise, I would like to be supportive of this paper. However, there are several aspects that I have problems with, or that I think are too simplistic. And in the final assessment I am not convinced they have proved the premise. I am not convinced that these issues can be resolved. As such I do not recommend publication.

A: We thank the reviewer for his/her time and effort and helpful comments. We agree in most aspects with the reviewer.
We feel, however, that we failed to convey the aim or premise of the paper which apparently has caused some confusion. We set out to illustrate that today's models include different (sometimes contradicting) approaches (or paradigms) which are not always obvious at first sight because they are "somewhat hidden" in a set of complex equations. This is problematic as even seemingly minor changes in individual (uncertain) model parameters (or formulations) can introduce considerable uncertainties in projections. We conclude that more observations are mandatory for the development of reliable models. To this end our results suggest that estimates of nitrogen fixation alone appear to be insufficient and propose to add in-situ observations of diazotroph biomass.

That said, we will state our aim and premises more clearly in a revised version of the manuscript. This may well include to change the title because, in hindsight, we feel that the word "juxtaposition" might have been misleading. Additionally, we will add more examples that illustrate that both paradigms are actively used in state-of-the-art models that support political decision making (please see Munkes, Löptien and Dietze, 2020, submitted: https://www.biogeosciences-discuss.net/bg-2020-151/ for an extensive list). Please note in this regard that our study differs substantially from Landolfi et al, 2015, who, back then lacked the resources to adjust model parameters systematically to match observational estimates (which is different to date because computers have become more powerful). Consequently, Landolfi et al. 2015 showcase very different

patterns of simulated nitrogen fixation when using differing model approaches (presumably because they were not able to optimize the respective models until "equally-good" fits to observations were reached).

Please note also that the purpose of this study is not to advocate one paradigm over the other and we agree with the reviewer's opinion that a single paradigm may well be too simplistic (and we will explicitly voice this in the revised version of the manuscript). Our scope here is to trigger a discussion on what it takes to realistically simulate the dynamics of diazotrophs. We aim to work towards reconciling the following opposing views in the field: (1) Numerical modelling of diazotroph dynamics is advanced and robust enough to be used for political decision making and (2) our understanding of diazotrophs as expressed in models is too simplistic and - as a consequence - flawed.

R: 1) The two paradigm concept is far too simplistic, and I believe unrealistic. See reviews by Sohm et al (2011) and more recent by Zehr and Carpone (2020), where much of the discussion of controlling mechanisms is focused on iron/phosphate availability perspective. In particular, the importance of iron is neglected in these paradigms and likely to be a more important than either grazing or phosphorus demands (see e.g. Ward et al 2013; Schlosser et al 2014). Moreover, I am not convinced that the "selective grazing" is a paradigm used by many models as stipulated (further references would be needed to show this to be true "de facto standard", line 247 needs substantiated). Early models of diazotrophy were based on Trichodesmium, which indeed appears to have lower grazing pressure and thus earlier models may have incorporated this type of parameterization. But it is now known that there is a great variety of diazotrophs (see e.g. Zehr et al 2020) and many do not appear to be grazed less than other phytoplankton. So this "paradigm" appears highly flawed. In fact, a study cited in this paper, Wang et al (2019) show a case where parameterizing reduced grazing on diazotrophs led to an unrealistic distribution of diazotrophs. It appears that such results are also found in this study (line 205-206). So why even make this a "paradigm"? Similarly, I am not convinced that the P-demand paradigm is fully justified. The study by Landolfi et al (2015) appears to have a very different parameterization of phosphate acquisition. It would seem that at least an iron paradigm should have been included (instead). Line 244: "...exploring two paradigms that are proposed in the literature" is too strong a statement. These do not appear to be the major paradigms that have been put forth (see reviews suggested above). Given the diversity of diazotrophs, it is likely that many processes lead to nitrogen fixation patterns, and expecting any single paradigm to explain them is simplistic. And as such, the setup of the paper appears fatally flawed.

A: We apologize for the confusion. The manuscript is not intended to promote any paradigm over another. We will make this clearer in the revised version of the manuscript. The ratio behind our choice of paradigms presented in this study is that they are currently put to use in applications, targeted to aid political decision making: paradigm one, coined "selective grazing" in our study, is e.g. at work in the models underlying Dzierzbicka-Głowacka et al. (2013), Keller et al. (2012), Paulsen et al. (2017) and Savchuck (2002). Paradigm two, coined "low P-demand" in our study, is at work in, e.g., in the SCOBI-model by assuming a lower half-saturation constant for P for nitrogen fixing cyanobacteria than for diatoms (Eilola et al., 2009) as well as in the BALTSEM-model which uses a lower half-saturation constant for cyanobacteria than for "summer species" (Savchuck, 2002).

We will add these extra model references along with supporting observational evidence that some diazotrophs have relatively low P-requirements (e.g., Degerholm et al., 2006) or might be able to adapt to low P-environments (e.g., Wu et al. 2011) to the revised version of the manuscript.

Ultimately we agree with the reviewer that the underlying logic of numerical models of diazotroph dynamics is often at odds with evidence from in situ and laboratory studies (mainly because very different species are summarized in a single functional group to keep the model complexity on a reasonable level). Please note that we started a painstaking process of reconciling this in Munkes, Löptien and Dietze, 2020 (submitted: https://www.biogeosciences-discuss.net/bg-2020-151/) for the Baltic Sea where data coverage is especially high. This paper is intended to illustrate the effect of certain model assumptions on the uncertainty of future projections. We agree that our configurations do not encompass the full range of respective uncertainties and will voice that more clearly in the revised version of the manuscript. Further we agree that our incomprehensive understanding of oceanic iron dynamics adds considerable uncertainty (and we will discuss this in the revised version of the manuscript). Among the problems associated with effects of iron dynamics on diazotrophs is that data sets of iron are so sparse that the current CMIP6 protocol states that the "... actual initialization" of Fe "... is left to the discretion of each modelling group ..." (Or et al, 2017, page 2175). Further, residence times of iron in contemporary models differ by two orders of magnitude (Tagliabue et al. 2016) which showcases that both, sources and sinks of iron, are not comprehensively understood. The combination of sparse information on standing stocks along with an incomprehensive understanding of sources and sinks provides a limit on

simulating the effect of iron on diazotrophs with the current generation of coupled ocean-circulation biogeochemical models.

That said, we have - encouraged by reviewer's criticism - performed additional control simulations where we raised the half saturation constant for iron for the diazotrophs by a factor of two, in order to test for robustness of our results. In these experiments C_d is increased by 0.1 for all simulations to obtain reasonable fixation rates. The results regarding our two paradigms are very similar to the results presented in our manuscript. The simulated biomass of diazotrophs is (as expected) slightly less in the Southern Hemisphere and slightly enhanced in the Atlantic and in the Indian Ocean. The respective future projections (RCP 8.5 scenario) are depicted below.

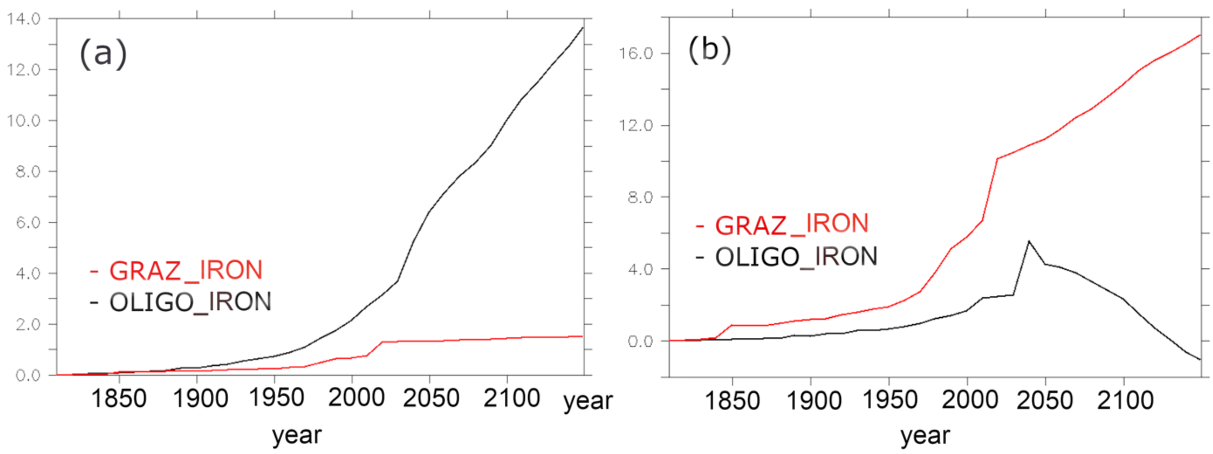

Fig.1: *Anomalous projected evolution of (a) global annual mean biomass of diazotrophic biomass in units Tg C and (b) global annual mean nitrogen fixation in units Tg N yr*$^{-1}$*. The model setups GRAZ_IRON and OLIGO_IRON are identical to the setups in our manuscript apart from doubling the half-saturation for iron for diazotrophs and increasing C_d by 0.1 for both simulations.*

Orr et al. (2017): Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP), Geosci. Model Dev., 10, 2169–2199, https://doi.org/10.5194/gmd-10-2169-2017

Tagliabue et al. "How well do global ocean biogeochemistry models simulate dissolved iron distributions?." *Global Biogeochemical Cycles* 30.2 (2016): 149-174.

We will add the information that our results are apparently robust towards changes of the half saturation constant for iron relative to the default settings used also in: Getzlaff and Oschlies, 2017; Getzlaff and Dietze, 2013; Keller et al., 2014; Kemena et al., 2019; Löptien and Dietze, 2017, 2019; Somes and Oschlies, 2015; Reith et al., 2016.

R: 2) Line 95: would have been better to be clearer what you mean by iron being not being explicitly resolved. Does that mean iron concentration are imposed? (I note that it is not a state variable). Given that iron is likely important in controlling diazotroph distributions, this suggests in itself that this is not the best model for exploring controls on diazotrophy.

A: We apologize for the confusion. The model indeed uses a fixed iron mask and no prognostic iron component. We will clarify this in the revised version of the manuscript.

3) All simulations (REF, OLIGO, GRAZ) have the assumption that diazotrophs do not grow above a temperature threshold and that they are dis-advantaged in nitrate-replete water (though this latter parameter is one that is explored, but only within a narrow range). Could it be that these two assumptions are responsible for the similarities between the simulations in the current day model ocean. That is: any other assumptions (as in OLIGO or GRAZ) are slave to these other very strong restrictions. And then that it is the expansion of warmer, lower nitrate conditions in future change simulation that allow the two simulations to diverge. Put another way, these other assumptions (temp, NO3 handicap) are stronger controllers of the nitrogen fixation. So a "bad" parameterization is constrained by the temperature/handicap assumptions in such a way that it doesn't show up until warming occurs. This does not totally detract from the premise of the paper, but it does suggest that a "bad" parameterization could lead to unrealistic future projections. Which is an important difference to the premise.

A: Yes, this is the point of the paper: a "bad" parameterization could lead to unrealistic model sensitivities to climate change. And yes: without a comprehensive understanding, any parameterization may well have thresholds which once reached may set loose unrealistic model behaviour. The idea of this manuscript is to illustrate that two apparently similar model approaches may well be founded on very different paradigms and - consequently - feature diverging projections. We feel that this is an important point because both model approaches presented in our study do already today influence political decision making (e.g., Meier et al., 2014: Ensemble Modeling of the Baltic Sea Ecosystem to Provide Scenarios for

Management, AMBIO). We will make that clearer in the revised version of the manuscript.

4) The above also leads to the question on how reasonable the temp/handicap parameterizations are? There are cold water diazotrophs – as suggested by Harding et al (2018), by diazotrophy in places such as the Baltic Sea, and as shown in the Wang et al (2019) estimates shown in Fig 1d? A modelling study (Monteiro et al 2011) has shown that temperature does not need to be invoked to explain diazotroph distribtions. By constraining diazotrophy by temperature you have forced it to be close to observations, but not necessarily for the right reasons. How necessary is the NO3-repletehandicap? I feel as though these two parameterizations should be far more fully understood before taking on this type of "paradigm" project.

A: We agree and will add a paragraph on temperature dependency in the revised version of the manuscript. Please note that the temperature dependence is still applied in many models (which does not mean that we necessarily think this is correct), even in Baltic Sea models: c.f. Munkes, Löptien and Dietze, 2020, submitted: https://www.biogeosciences-discuss.net/bg-2020-151/, their Fig.1.

As far as we know, the functional relationships between diazotroph growth and temperature used in today's models does typically not provide diazotrophs a real advantage over ordinary (non-fixing) phytoplankton in the sense that even under optimal temperature conditions, the growth rates for diazotrophs are typically slower than for ordinary phytoplankton. This leads to a situation where diazotrophs - ultimately – destroy their own ecological niche by supplying nitrogen to faster-growing ordinary phytoplankton with which they compete for other resources (phosphate, iron, light).

We apologize for being unclear in our manuscript in this respect: we did not explicitly formulate a NO3-repletehandicap. The NO3-repletehandicap is a consequence of the relatively slow growth of diazotrophs. If there is enough N, P and light, and if the zooplankton is not grazing selectively than in resource competition we will see fast-growing ordinary phytoplankton as a winner and the "looser" are the relatively slow-growing diazotrophs.

R: 5) Using Wang et al (2019) for the skill assessment seems awkward since Wang et al (2019) is itself a model estimate. Though data constrained, I would suggest it is not a good benchmark. Deutch et al (2007) was also a "data constrained" estimate and it is very different to that found in Wang et al (2019). Wang et al (2019) is far more believable and a better study, but this example *does suggest that there remains significant level of uncertainty even in a data constrained model,*

A: We agree with the reviewer that the fact that we had to use a model-aided data-constrained estimate of nitrogen fixation rather than a climatological estimate based solely on in-situ measurements of fixation rates is indicative of an awkward situation, characterized by a lack of observational data. In fact, this is one of the major messages of our manuscript: a prerequisite for reliable projections of the fertilizing effect of diazotrophs are additional observational data. Our manuscript illustrates that even a climatology comparable to the Wang et al (2019) model-aided estimate in terms of spatial coverage may not suffice to dissect major controls of diazotroph dynamics because very different model approaches may well result in equally-well fits to such a climatology. Thus, nitrogen fixation estimates should ideally be complemented with a yet to be sampled climatology of in-situ observations of diazotroph biomass (which differs strongly between our model versions). We will make this point more clear in the revised manuscript.

Unfortunately, we are a long way from coming even close to an in-situ data density comparable to Wang et al (2019). This is why, for the time being, we rate the Wang et al (2019) as arguably the most recent and comprehensive estimate of global pelagic nitrogen fixation. We will add a respective discussion to the revised version of the manuscript. This discussion will elaborate on the fundamental difference between our (forward) models and the data-constrained (somewhat inverse) method of Wang et al (2019) so that it becomes clearer why it makes sense to compare "models" with "models".

R: 6) Why not show the future diazotroph/nitrogen fixation distributions? Does GRAZ become unrealistic? I felt that since this was the crux of the premise, this last part of the paper was very rushed through: paragraphs only and one very simple figure. There is a mention of Bay of Biscayne feature (line 240), but this is not shown and appears rather arbitrary.

A: Good point. Agreed! We will include these figures in the revised version of our manuscript. In our model the overall pattern of the biomass of diazotrophs shows little changes relative to today (apart from the Indian Ocean, as described in line 240ff. - cf., Fig 2 below).

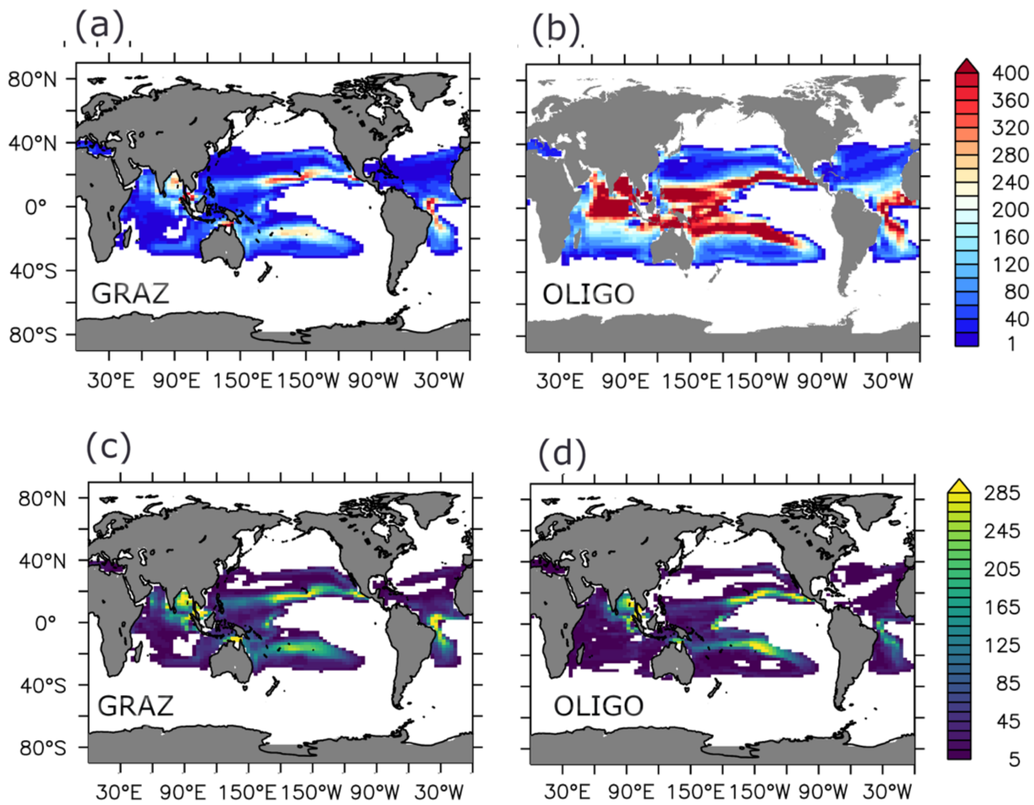

Fig.2 (a, b) *Projected annual mean non-zero diazotrophic biomass for the year 2100 integrated over the upper 100m in mg C/m^2.* ***(c.d)*** *Simulated nitrogen fixation in mmol N/m^2/yr in the year 2100.*

<u>Details:</u>

- R: Line 47: By "supply" do you mean "concentrations"? I would agree that concentrations do not necessarily correlate – but do not think that studies have shown the "supply" doesn't correlate as it is so difficult to measure supply rates.

- A: Thanks – we will replace "supply" with "concentrations"

- R: Line 89: Do you mean "DIC" not "DIN"

- A: Yes, thanks!

R: Line 191: "desert dwellers" does not seem appropriate term here

A: O.K.

R: Line 193/256: The use of the word "quota" does not seem right here. Do you mean "ratio" instead?

A: Ratio might indeed be the better expression.

R. Line 194: I do not understand what you mean here? How do they go below zero?

A: The discretization of numerical ocean models inevitably induces spurious effects on ocean transports (e.g., Hofmann, M., & Morales Maqueda, M. A. (2006): Performance of a second-order moments advection scheme in an ocean general circulation model. *Journal of Geophysical Research: Oceans*, *111*(C5)).

All numerical approaches we know of are situated within a triangular with the following edges: high computational cost, spurious diffusive behavior and spurious dispersive behavior. The spurious dispersive behavior can cause (very small) negative concentrations. There are various ways to mask these. We opted for mentioning them.

R: Line 218: why do discuss only phosphate here. The changes to nitrate are also important to the issue under discussion.

A: Agreed, we will add a discussion on nitrogen.

R: References:

- *Deutsch et al (2007). Spatial coupling of nitrogen inputs and losses in the ocean. Nature, 445, 163-167.*
- *Harding et al., Symbiotic unicellular cyanobacteria fix nitrogen in the Arctic Ocean. Proc. Natl. Acad. Sci. U.S.A. 115, 13371–13375 (2018). doi: 10.1073/pnas.1813658115;*
- *Landolfi, A., W. Koeve, H. Dietze, P. Kähler, and A. Oschlies (2015), A new perspective on environmental controls of marine nitrogen fixation, Geophys. Res. Lett., 42, 4482–4489, doi:10.1002/2015GL063756*
- *Monteiro et al (2011), Biogeographical controls on marine nitrogen fixers, Global Biogeochem. Cycles, 25, GB2003.*
- *Schlosser, C., et al. (2014), Seasonal ITCZ migration dynamically controls the location of the (sub)tropical Atlantic biogeochemical divide, PNAS, 111(4), 1438–1442, doi:10.1073/pnas.1318670111.*
- *Sohm, J.A., E. A. Webb, D. G. Capone, Emerging patterns of marine nitrogen fixation. Nat. Rev. Microbiol. 9, 499–508 (2011). doi: 10.1038/nrmicro2594*
- *Wang, W.-L., J. K. Moore, A. C. Martiny, F. W. Primeau, Convergent estimates of marine nitrogen fixation. Nature 566, 205–211 (2019). doi: 10.1038/s41586-019-0911-2*
- *Ward, B. A., S. Dutkiewicz, M. Moore, and M. J. Follows (2013), Iron, phosphorus, and nitrogen supply ratios define the biogeography of*

nitrogen fixation, Limnol. Oceanogr. 58(6), 2059–2075, doi:10.4319/lo.2013.58.6.2059.

- *Zehr and Capone (2020). Changing perspectives in marine nitrogen fixation. Science, 368 eaay9514*

A: We thank the reviewer for the suggested references and will include those which have not been listed already in the revised version of the manuscript.