Response to comments by both reviewers for "Data-based assessment of environmental controls on global marine nitrogen fixation" by Luo et al.

Y.-W. Luo, I. D. Lima, D. M. Karl, C. Deutsch and S. C. Doney

First of all, we'd like to thank the anonymous reviewer #1 and Dr. Gruber for their detailed and constructive comments that will greatly improve the quality and clarity of our manuscript.

After carefully study of the reviewers' comments, we feel that the contribution of this manuscript to the community can be a first-stage analysis of the recently compiled N_2 fixation dataset. We will emphasize what current hypotheses of environmental controls on marine N_2 fixation are supported by the data, and what hypotheses are questionable based on current data availability and need further field investigation.

The following are responses to two major comments that both reviewers have raised:

(1) The current evidence does not exclusively support our conclusion that regional N loss in form of subsurface denitrification is one of the main environmental controls on N_2 fixation. We agree with the reviewers: As pointed by the reviewers, the N_2 fixation rates only quickly drop after the minimum oxygen level is higher than 150 µM, which does not necessarily indicate denitrification. Also, the subsurface denitrification may not directly connected to its overlying surface waters but is more likely to influence the downstream areas. Thus we will take out this conclusion from the abstract and the main text, and instead only state " (why) the subsurface minimum oxygen level is one of the currently available best predictors for the observed spatial pattern of the N_2 fixation". We will further discuss (and also raise the question) what are possible mechanisms that higher subsurface minimum oxygen can limit regional N_2 fixation, such as the last paragraph of Section 4.1 in the original manuscript.

(2) Both reviewers suggest us to use the modeled P* convergence (i.e. flux) instead of P* concentration as one of the environmental controls, which can be used to further investigate if the denitrification controls N_2 fixation. We have obtained the N_2 fixation rates diagnosed from modeled P* convergence from Dr. Curtis Deutsch (who will also be added as a new co-author). These are same data shown in Deutsch et al. 2007 (Nature) with P* convergence of both inorganic and organic matter considered. They are also approximately log-normal distributed. However, the R^2 values of the linear regression and the quadratic regression between the P*-convergence-diagnosed and the observed N_2 fixation are low, only 0.07 and 0.11, respectively. This supports that the current available data cannot prove that the denitrification is a major driver for N_2 fixation. Figure C1 is the full comparison between the P*-convergence-diagnosed and the observed N_2 fixation: the N_2 fixation rates diagnosed from P* convergence are largely comparable to the observations in the Pacific, capture the high N_2 fixation in equatorial Atlantic and low N_2 fixation in the North Atlantic Subtropical Gyre, but overestimate in the South Atlantic. We will add Figure C1 and the related discussion in the revised manuscript.

Responses to other comments from Reviewer #1:

(1) Reviewer #1 suggest to emphasize caveats of different types (fluxes vs. standing stocks) of environmental parameters before they are presented. Reviewer #1 also raise an example that the standing stocks of nutrients may correlated or anti-correlated to phytoplankton biomass during different stages of their life cycle.

Response: We agree and will highlight this limitation in Section 2.1 Data sources. In fact, for the environmental parameters we have tested, only nitrate, phosphate, and the P* concentration derived from the former two, are standing stocks. SST is not a flux, but is highly correlated to solar radiation. As stated above, after information of P* convergence, which is a flux, is included in the manuscript, this weakness regarding nutrient standing stocks not fluxes can be partially addressed.

(2) For the same reason, Reviewer #1 also suggest us to do same regression analysis for the surface volumetric N_2 fixation rates, as some of the environmental parameters only represents surface standing stocks.

Response: To clarify, some of our depth-integrated data were reported directly by original data contributors (they sampled whole water column) and the others are calculated by us if the original data contributors reported a vertical profile. In the latter case, we practically assumed the scientists in the fields knew the deepest depth where significant N_2 fixation existed and sampled down to that depth. There is another separate dataset that contains all the volumetric N_2 fixation rates. Thus the depth-integrated and the volumetric N_2 fixation rates in the database each have their own data sources while overlap in those data points belonging to a vertical profile. Please refer to our ESSD paper (Luo et al, 2012, doi:10.5194/essd-4-47-2012) for full description of the database.

We took out the volumetric N_2 fixation rates in the surface 25 m and did same regression analysis with the environmental parameters. The correlation, and the R^2 values in both simple regression and multiple linear regression (MLR) are much lower than those when the depth-integrated rates are used. For instance, the MLR with all the environmental parameters included as predictors only reached a R^2 of 0.30. One possible reason is that the high variance of volumetric N_2 fixation may exist with depth, even in the surface 25 m. This certainly deserves more analysis in the future. But we choose not to include these volumetric data, as in this paper we aim to do only a first-stage analysis. Again, with P* convergence data included, all the nutrient data point out that the current available N_2 fixation data are not enough to prove or reject the related hypotheses.

(3) Reviewer #1 states: use of a quadratic to describe the response of (log transformed) variables could be considered arbitrary. Significantly, this occasionally resulted in fitted responses being non-monotonic, which surely means that the performed analysis sometimes contradicts the expected responses/hypothesis? e.g. see Figs. 3 b & d? The authors acknowledge this in the case of dust deposition (Page 7378), but effectively

dismiss it in the case of thermocline O2 (Page 7380), why should these be treated differently?
Response: We agree that the quadratic regression with log-transformed variables could be arbitrary, but we believe it is appropriate in this first-stage analysis (another reviewer, Dr. Gruber, specifically raised this issue in his comment and also agreed that the method used in the paper is OK at this stage). Reviewer #1 is correct that the performed analysis sometimes contradicts the expected response and the fitted quadratic curves can be non-monotonic. Actually we looked at the whole range of the variable range and checked if in the majority of the range if the curve is contradicts to the expected response or not. This has been specifically address in Table 3 and the related texts.

(4) Reviewer #1 states: Regarding the modelled dust flux, the authors could also have used modelled soluble atmospheric iron fluxes as well as potentially providing comparisons with other atmospheric dust models. Additionally, I wonder whether model based estimates of other potential Fe sources could also be included. Irrespectively, as the Fe related variable is based on a model product, some sensitivity analysis would certainly seem appropriate.

Responses: current global dust deposition data are all based on models and only very limited observations are available. The distribution of soluble Fe is also mainly based on models. We admit that the modeled dust deposition products we used have many limitation, but at least it captures the basic pattern of the dust deposition that most scientists would agree. In addition, the modeled dust deposition rates range on several orders of magnitude and are approximately log-normal distributed. As we are using log-transformed dust deposition in the paper, the sensitivity of the outputs from the dust deposition model may not influence the analysis greatly. We did not do the dust deposition model but only took the outputs to be used in our analysis. Sensitivity test for the dust deposition model is beyond the scope of this paper.

Anyway, as discussed in the manuscript, the data show that the Pacific tends to have higher N_2 fixation rates than the Atlantic in most regions, while we all know dust deposition in the Atlantic is much higher than the Pacific. Even different dust deposition products are used, this conclusion is not likely being changed.

(5) Figure 1b. Did the authors do a test for log normality?

Response: it does not pass the log normality test, mostly because it is not symmetrical. We only justify from the figures that they are "approximately" log-normally distributed, at least the distribution is more likely in a "bell" shape after log-transformed.

(6) Fig. 3. Why are plots not included for the other environmental variables?

Response: we did not include plots for other environmental variables to save space and also think those were not key variables. We will include all the plots in the revised manuscript.

(7) Fig. 6. It wasn't immediately obvious how to compare these colour scales. Can the same scales be used for both sub-panels?

Response: We will change the lower panel (error) to the same unit and scale as the upper panel.

We'll also revise the manuscript according to the other minor specific comments by Reviewer #1.

Responses to other comments from Reviewer Dr. Gruber:

(1) Dr. Gruber suggests us to tighten up the manuscript, and also remove the texts describing the various choices we considered. Dr. Gruber also suggests to remove wind entirely from the paper as a predictor.

Response: We will go over the paper, following Dr. Gruber's suggestions, and make the paper more concise. By remove all the related contents related to wind, the paper should be shorten further.

(2) Dr. Gruber asks if the statistical model can be implemented to $^{15}N_2$-based and C_2H_2-based nitrogen fixation measurements separately.

Response: (This issue actually has been raised and somehow discussed in the earlier ESSD database paper.) Most of the data points were measured using the $^{15}N_2$ method. Although number of points from the C_2H_2 are fewer, they are in some key locations such as in the equatorial and the south Pacific, where the samples from the $^{15}N_2$ method are very scarce. Another reason we chose to include measurements from both methods in one analysis is that the N_2 fixation rates were log-transformed in our analysis. The difference between the two methods after the log-transformation may not influence our linear analysis greatly.

(3) Dr. Gruber suggests to add mean solar radiation in the mixed layer as a predictor.

Response: We have tried the mean solar radiation in the mixed layer. It was calculated by using light attenuation coefficient of 0.045 (which leads to a bottom euphotic zone depth of 100 m). It was slightly better correlated to N_2 fixation than the mixed layer depth, but much less than the surface solar radiation. The R^2 values of quadratic regressions using surface solar radiation, mean solar radiation in the mixed layer, and the mixed layer depth were 0.38, 0.23 and 0.19, respectively. When the mean solar radiation in the mixed layer is included in the multiple linear regression, it replaced the mixed layer depth, and the R^2 value slightly improved from 0.58 to 0.59. We will include mean solar radiation in the mixed layer in the revision, with a caution that it partly depends on the light attenuation coefficient selected.

(4) Dr. Gruber also has the same suggestion as Reviewer #1 for the different scales of the two panels in Figure 6.

Response: We'll change the error panel to the same scale of the estimates.

(5) Dr. Gruber comments that the low-latitude phytoplankton tend to have higher N:P ratios may impact the P* convergence results.

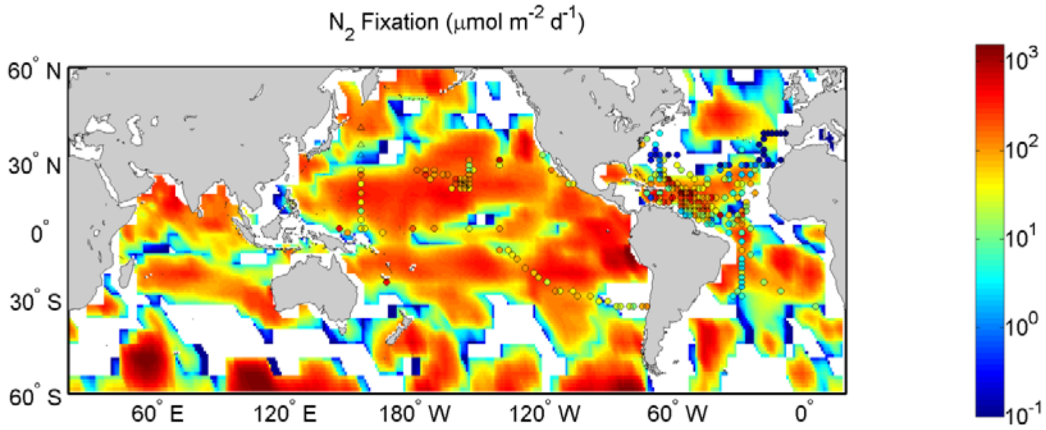
Response: We'll include this issue in the discussion for the P* convergence section in the revised manuscript.

(6) Dr. Gruber requests to explain how the anomalies are calculated in Fig. 7.

Response: The are spatial anomalies of each component in the MLR equation. We included the explanation in the last paragraph of Section 2.3. Apparently it was not clear enough and we'll clarify in the revision.

We'll also revise the manuscript according to the other minor specific comments by Dr. Gruber.

(a)



(b)

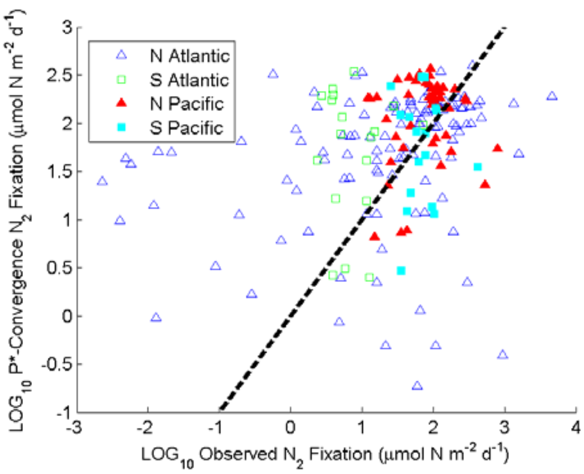


Figure C1. Comparison between the observed N_2 fixation rates (binned on 1˚×1˚) and the P*-convergence-diagnosed N_2 fixation rates. (a) P*-convergence-diagnosed N_2 fixation rates in global ocean (background) and the observed N_2 fixation rates (filled circles) on same color scale. The white background represents area of zero P*-convergence-diagnosed N_2 fixation rates. Several empty triangles represents observed zero N_2 fixation rates. (b) Comparison of the two datasets on same locations with different markers for different basins. The dashed line is the 1:1 line.