Response to Referee #4

Below are copied the comments in regular font with our point by point responses in blue. Changes in the manuscript appear in 'track change' mode.

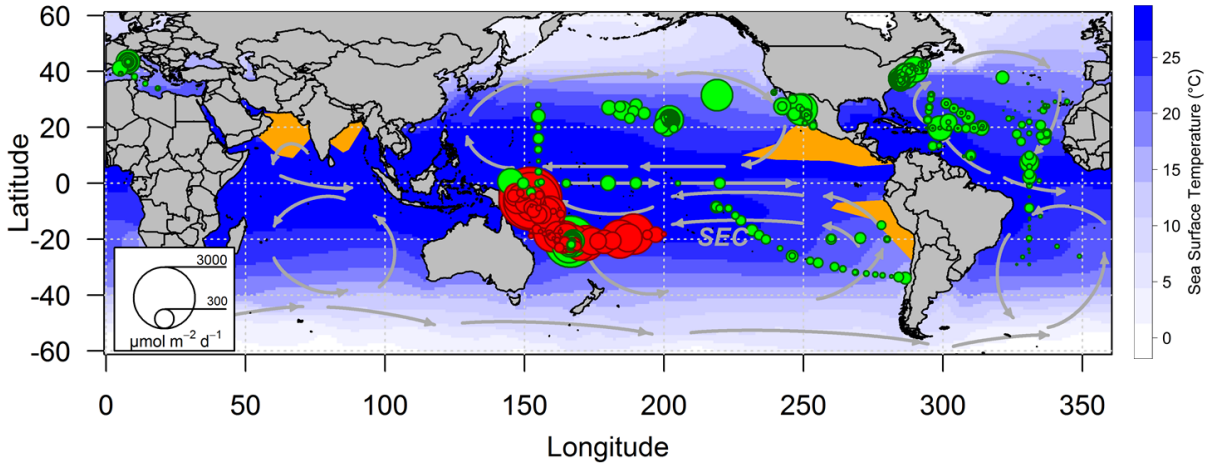
Review of the manuscript: In depth characterization of diazotrophs activity across the Western Tropical South Pacific hot spot of N2 fixation by S. Bonnet et al. I can't recommend this manuscript for publication in its current form, which is a shame, because it partially presents a very interesting data set about a large, but relatively understudied part of the world ocean. I have made this decision primarily because the authors do not present sufficient data or a convincing analysis to make their case. I think they need to show the reader a lot more information, a more expansive and rigorous analysis and be more up front about how they arrive at their conclusions. At present, I am unconvinced. I don't think that the problem can be simply rectified by rejigging the existing manuscript, adding some additional figures/tables or doing more statistics as in a normal 'major' revision. The authors need to make a fresh start, think carefully about where they want to go and the strengths/limitations of their data, and then re-write carefully. If they do that properly, I think this work has the potential to make a very significant contribution to the global N-fixation literature.
We are quite surprised that the reviewer opinion is that this paper do not present sufficient data. However, we believe that our story is compelling because:
-we provide accurate bulk N2 fixation over a 4000 km transect and the whole photic layer, which represent the first data for this region
-we tested several methods of incubations and provide a critic opinion of each method in the context of the methodological issues associated with N2 fixation measurements in the N2 fixation community
-we further investigate the contribution of the main diazotroph groups using single-cell isotope approaches (nanoSIMS)
-we provide extensive environmental data (dissolved iron, DIP turnover times…) to discuss the spatial variability of the N2 fixation activity observed. This part has been improved following the suggestions (please see below)

Some general comments: This reviewer recognizes that English may not be the lead author's prime language. However, in places, some of the word choices and syntax are ambiguous so that different readers are likely to take different things away from the manuscript. Specific examples are noted below. The authors should take advantage of the native English speakers in the author list to ensure that the wording is done with more precision to ensure that the intended meaning is communicated.
We acknowledge that English was not perfect as it is not the native language of the first author. The revised version of the manuscript has been checked by the native English speaker of the author list.

Repeated references are made throughout to works which are 'submitted' or 'in review'. While this indicates the present manuscript is quite timely, the reader can make no critical or objective use of these references as they haven't seen the light of day, might be rejected or heavily altered before they are eventually published. It's hard to take these citations at face value. Where possible, I would include actual data from closely related studies in your paper to genuinely demonstrate your point, noting by citation that a fuller description will be published in other work. If not possible, stick to your own dataset.
We totally understand this comment. The OUTPACE special issue is divided in 27 papers, each treating a specific part of the 'story'. Since the submission of this manuscript, several of the cited papers are now available online on the special issue webpage https://www.biogeosciences.net/special_issue894.html, and some are accepted (this has been updated in the revised version), which should improve the accessibility.

The authors need to be a lot more precise about your geography. The 'regions' used herein are quite large, loosely defined and contain a number of somewhat similar, but different oceanographic regimes. For example, 'Western Tropical South Pacific' would also include the Warm Pool region north of PNG – it's western, south of the equator and certainly tropical – but a different setting altogether. Likewise, 'Eastern Tropical South Pacific' also includes extensive areas of Ekman driven upwelling where it would be difficult to extrapolate your measurements. The longitude scale on Fig. 1 (bottom) is seriously wrong.
We agree that the WTSP is a vast region that includes the warm pool region north of PNG, but also the Solomon Sea, the Coral Sea etc… these regions were previously documented for N2 fixation, and we clearly state in the introduction section that our goal is to specifically study the central and eastern parts of the WTSP, that are critically undersampled + the border of the south pacific gyre. This was the goal of the OUTPACE project, so we only report results from this cruise. To avoid any misunderstanding, we have added 'OUTPACE cruise' in the title of the manuscript to specify that this article is about the results of this specific cruise and not the results about the entire WTSP.
The longitude scale on Figure here is the conventional one decided for all outpace papers of the Special issue. However we fixed that point in the revised version.

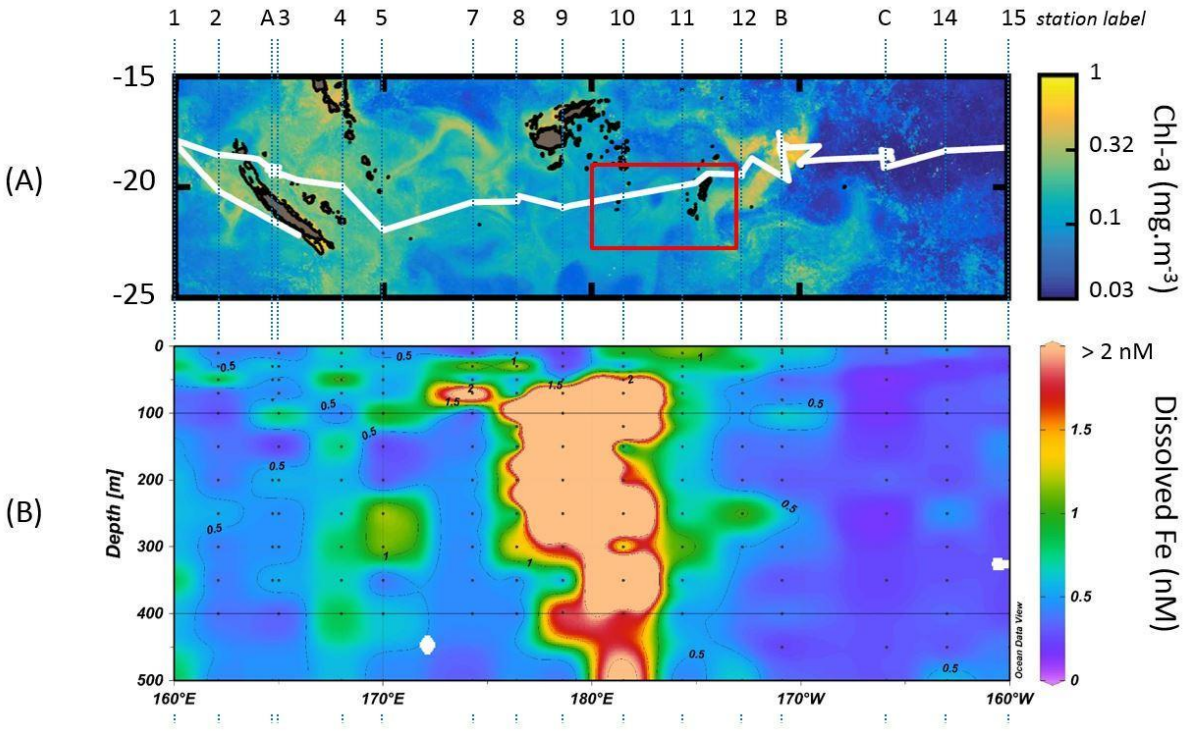
For starters, what exactly do you mean by a N-fixation 'hot spot'? What's the cut-off? A summary table at least summarizing ranges of measured or reliably estimated N-fixation in other 'hot-spot' parts of the world ocean (e.g. Arabian Sea, Caribbean, Arafura Sea, etc., etc.) would be very useful to set the scene and would tie your work into the wider literature of global N-fixation. As a reader, I'm thinking a lot about how is this paper compares with the much larger body of published work done by Capone, Carpenter and their collaborators/students, etc. There seems to be little quantitative integration (show me the numbers) with even the many more recent measurements of N-fixation in the SW Pacific. These need to be tied together, or shown why not.
To the best of our knowledge, there is no 'official' cut-off to design a hot spot but we clearly mentioned in the manuscript that the rates reported here are in the upper range of the higher caregory (100-1000 µmol N m^{-2} d^{-1}) of rates reported in the N_2 fixation MAREDAT database for the global ocean (Luo et al., 2012). This was our argument to say that it is hot spot. However, as suggested by Reviewer 1, we tone down the hot spot story throughout the revised version. We also recently published the map below, which gathers N2 fixation rates from the Luo et al. global database (in green) compared to the rates measured in the WTSP (in red, this cruise and others) by our team. It speaks for itself and reveals the WTSP as a high N2 fixation area in the global ocean, at least from our current knowledge. We acknowledge that there are also other high N2 fixation areas such as the Caribbean Sea and probably other parts which are to date particularly undersampled. This map was published in a very short letter in PNAS to argue for a spatial decoupling between N2 fixation in the western Pacific and N losses in the eastern Pacific, and only present depth-integrated rates, with no details. The present paper aims at describing vertical and horizontal N2 fixation rates during the OUTPACE cruise in relation with environmental parameters, identify the main players etc…



From Bonnet, S., Caffin, M., Berthelot, H., and Moutin, T.: Hot spot of N2 fixation in the western tropical South Pacific pleads for a spatial decoupling between N2 fixation and denitrification, Proceedings of the National Academy of Sciences of the United States of America, 114, E2800-E2801, 10.1073/pnas.1619514114, 2017.

Figures and tables – Need more of them!, and more quantitative! ODV color contour plots are nice, but awfully hard to interpret quantitatively, and quite impossible if looking at a B/W version of your paper. Show some convincing/representative profiles, hard contour lines and East-West quantitative values. The discussion, by and large, is mostly hand-waving and speculation. While quite a few papers are cited (several of which haven't been published), there is a lack of quantitative information and data presented from these studies with which to compare the authors' results, assess their veracity and draw comparisons. The bit about regional differences being due to iron (etc.) inputs from submarine volcanos is wholly speculative on the information provided. Not a shred of quantitative information is presented to back this assertion up. Might the regional difference in fixation be due to regional differences in wind stress and water masses which affects the depth of the mixed layer and vertical mixing through the thermocline? I'd like to see a more focused discussion.

We now provide the vertical profiles of N2 fixation in the supplementary information. The vertical and longitudinal variability of N2 fixation are now better described in the results section and discussed in the section 4. The depth of the mixed layer has been also added to the revised manuscript and is homogeneous across the zonal transect and may not explain the differences observed (all this is now presented and discussed). The full set of dissolved iron data (see figure below) cannot be provided in the revised version as they are used in another paper under (minor) revision in Scientific reports. However, the discussion on the role of iron has been updated page 14 lines 9-34 to give more quantitative data.



Surface Chlorophyll-a concentration (mg m^{-3}) during the 45-day transect of the OUTPACE cruise (A) (The ocean color satellite products are produced by CLS. Figure courtesy of A. De Verneil). (B) Cross-section of dissolved Fe nM (0-500 m). From Guieu et al., (Under review, minor revisions)

Why no comparison with the extensive work done on N-fixation, fluxes and driving processes done at station ALOHA? An opportunity is missed.
We performed an extensive comparison between the OUTPACE results and the ALOHA station data in our other OUTPACE paper by Caffin et al., (2018) regarding the functioning of the ecosystem, the role of N2 fixation on export etc…

Some specific comments:
Page Line(s) Comment 1 2 What exactly do you mean by 'hot spot' – see above 3
Please see comment above

10-11 By 'highest rates of N2 fixation' do you mean on an area-specific basis (I doubt it) or aggregate fixation on a regional basis primarily because of the very large area involved? The oxygen deficient zones of the eastern Pacific are due to higher regional productivity arising from Ekman-driven upwelling at the basin scale, not N-fixation. Indeed, fixation tends to be lower in upwelling regions. Simple N:P ratios are a poor predictor in this case.
Yes, I mean that Deutsch et al., (2007) predicted the highest rates of N2 fixation in the global ocean in the ETSP. I totally agree that this is a high productivity area mainly due to the upwelling, so it was counter intuitive to find N2 fixation there, but their argument was that the decrease of P* in this region would be the result of intense N2 fixation. This Deutsch et al study motivated many cruises in the ETSP to quantify N2 fixation there. Most of them found low rates (average range ~0-60 µmol N m^{-2} d^{-1}, (Dekaezemacker et al., 2013; Fernandez et al., 2011; Fernandez et al., 2015; Knapp et al., 2016; Loescher et al., 2014)).

10-20 This paragraph seems to do a logical U-turn 21 etc. References to regional N-fixation ignore large database of published historical estimates by Capone, Carpenter, etc., etc., etc.

I am not sure to get this comment. In this paragraph we review the existing literature in the ETSP and WTSP and all along the manuscript we compare our results with the global MAREDAT database, which includes historical data from Capone, Carpenter... There are very few studies in this South Pacific. I wish I could cite more papers including Capone, Carpenter... I may have missed something, so if you know more papers published for this area, please provide the references. Some of them include Capone results especially the Knapp et al. 2016 and Dekaezemacker et al. for which D.G Capone is co-author (he was also the chief scientist onboard the cruise from which these papers originate). The Montoya et al., 2004 paper in the Arafura Sea also include D.G Capone data…

4 11 Change "equalled to" to "reliably extrapolated to estimates of"
This has been changed

18 Change "previously undocumented" to "new"
This has been changed

22 What is the basis of 'selection' – suggest leave out
nanoSIMS analyses are very expensive and time consuming. It was this not possible to perform such analyses at all stations, and we thus selected three stations to perform those analyses based on the diazotroph community composition assessed microscopically onboard. We remove the terms 'selected stations' in the introduction and specify in the Methods section that those analyses were performed at only 3 stations

22 What do you mean by 'potential ecological impact' of N-fixation? Poor wording.
We changed by 'biogeochemical impact'

29 Suggest changing 'contrasted' to something about a gradient of conditions. What's the essential difference between 'oligotrophic' and 'ultra-oligotrophic'? Strictly speaking that's like saying something is 'more better'.
The depth of the deep chlorophyll maximum was the main criteria. We changed the sentence as follows: 'It covered a trophic gradient from oligotrophy (deep chlorophyll maximum (DCM) located at ~80-100 m) in MA waters around New Caledonia, Vanuatu, Fiji up to Tonga, to ultra-oligotrophy (DCM located at 115-150 m) in GY waters located at the western boundary of the South Pacific Gyre (see Introduction article Moutin et al., 2017 for details on the cruise implementation)'.

33 (etc.) By 'fluorescence', I presume you mean 'chlorophyll fluorescence' – say it because lots of other things fluoresce if you measure it right.
It has been changed

5 4 Use 'stored' instead of 'preserved. More importantly for low-level nutrient analyses – how long were the samples stored before analysis (hours, days, months)?
This has been changed. The nutrient samples were stored for ~3 months before analysis (the time for the 4°C container to be back from French Polynesia (Tahiti) to Marseille (France). Previous experience from our team have confirmed that this way of conservation ($HgCl_2$, 4°C) is a valid methodology.

7-10 Are there any actually published papers that describe these methods? Preferable. In the case of dissolved iron, the more widely used and less confusing notation would be: Fediss.
The manuscripts mentioned in the submitted version are now all available online on the webpage of the special issue https://www.biogeosciences.net/special_issue894.html and most of them have been accepted, except the Fe paper (presently in minor revision) , so the method for dissolved Fe concentration determination is now given in this section.

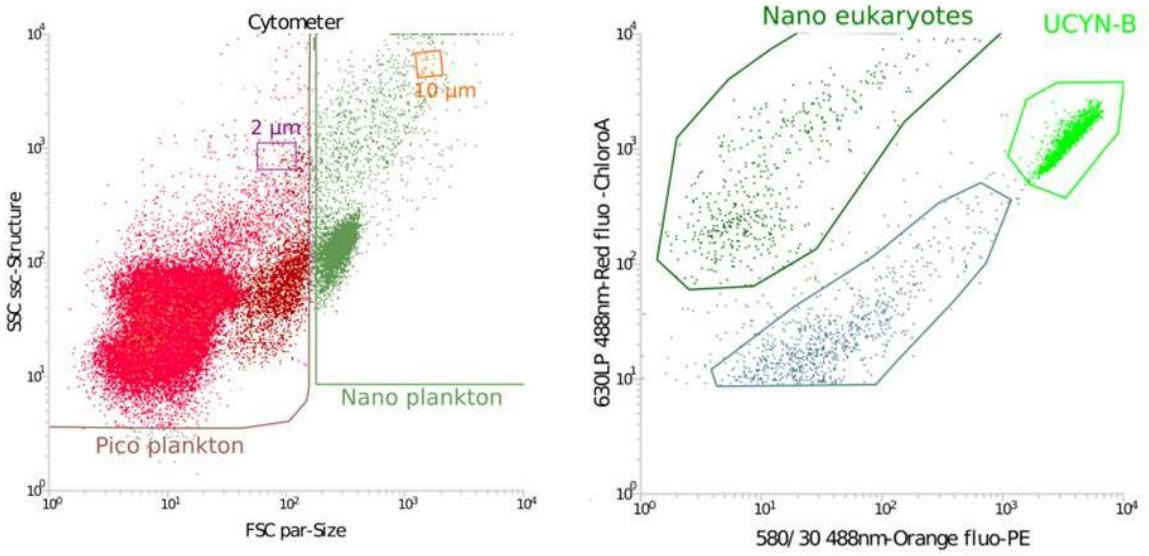
14 In using open-ocean communities, it is almost universally observed that metal and organic contamination results in under-estimation of rates due to toxicity of these materials to finicky oceanic bugs. Why do you think they are over-estimated?
The most likely metal contamination onboard a metallic ship is iron, which would likely fertilize the studied water mass and consequently would over estimate rates (together with DOM contaminations) (see previous work from Bonnet et al., 2009). We acknowledge that other metals such as Hg or Cu at high concentrations would have the opposite effect due to toxicity, but this is unlikely to occur with the protocols we use as they are scarce compared to iron.

20-21 Presumably you mean sub-samples taken from the Niskins. Explicit reading suggests you collected the water in situ in the poly-carb bottles.
This was stated in the first paragraph of the method section (common to all other sub-sections of the methods) 'Seawater samples were collected by 12-L Niskin bottles mounted on the CTD rosette'. Wee re-specified that seawater was collected from niskin bottles here.

33 How were these filters stored and for how long? Text suggests they were analysed almost right-away (good if you can do it correctly!).
We stored the dried filters in a desiccator for 3-4 months before analysis. As long as there is no water anymore on the filters, they can be stored for months.

6 20 What's the "them"?
The polycarbonate bottles mentioned in the previous sentence. We changed as follows for more clarity: 'eight additional polycarbonate (2.3 L) bottles were collected from the surface (50 % light irradiance) to determine *Trichodesmium* and UCYN-B specific N_2 fixation rates by nanoSIMS and quantify their contribution to bulk N_2 fixation. Two of these bottles were amended with $^{15}N_2$ as described above for further…'

26-28 What are the flow cytometric characteristics you sorted and counted the UCYN cells with? A lot of those don't have much, if any photosynthetic pigment, and if they did, the near-surface ones would likely be bleached a bit? It would be nice to see a cytogram.
Pico- and nano- phytoplankton were clustered on FSC vs. SSC cytograms (left panel on the figure below) using 2 µm beads. UCYNs were associated to the nano- phytoplankton (right panel on the figure) and the UCYN associated cluster was established using orange fluorescence vs. red fluorescence cytograms. The cluster corresponding to UCYNs is shown in light green on the cytogram of the right panel on the figure, and could be clearly separated from other clusters.



8 2 Strictly speaking, you're 'estimating' this, not determining it.
Yes, this has been changed

16 (etc) By 0-50 m, I presume you mean the 'surface mixed layer'. This is a key matter herein as the surface mixed layer thickness changes along your transect. Strictly speaking, surface is surface (say 0-5 m). Even small and ephemeral density gradients in the near-surface layer and surface mixed layer can have profound influences on vertical mixing rates and hence the light histories of cells embedded in the surface layer. Be very specific!
We changed the text to be more specific. 'The mixed layer depth (MLD) calculated according to the de Boyer-Montegut et al., (2004) method was located around 20-40 m throughout the zonal transect: Maximum temperatures were measured in the surface mixed layer (~0-20/40 m) and remained almost constant along the longitudinal transect with 29.1 ± 0.3°C in MA waters and 29.5 ± 0.4°C in GY waters'.

.

23 By DCM, I presume you mean 'deep chlorophyll maximum'?
Yes, it is now defined in the methods section.

29-34 It's not clear what, if anything, this paragraph contributes to the paper.
This paragraph is about a methodological aspect regarding the comparison of the solubilization of the 15N2 tracer according to the mode of incubation of samples used. This is the first time that such comparisons are done, and these results will be helpful for people who measure N2 fixation in on-deck incubators (a method widely used) and will hopefully convince them to perform MIMS measurements on their samples.

37-38 This seems very wrong. I'm presuming you actually mean the per-mil deviation of the particulate matter (δ15N) from the normal natural abundance of 15N (0.367%). Normally, N-fixation has a δ15N value close to 0‰. Larger deviations would suggest other fractionation processes such as denitrification. Clarify and fix up.
The term 15N/14N ratio was probably misleading. We changed the sentence by 'The natural N isotopic signature of suspended particles measured over the photic layer was on average -0.41‰ in MA waters and 8.06‰ in GY waters (Table 1)'.

9 1-8 This is all the 'new' N-fixation data in the results text. Pretty thin. As an interested party, I'd like to see a lot more. Graphics and tables too. 10-17 Correlations – So what?

Tables of correlation coefficients fill space, but are instantly forgettable. What's the point of the correlations other than that you can do them?
The presentation of N2 fixation rate results have been extended, we have merged sections 3.3 (N2 fixation results) and 3.4 (correlations) in the new version on the manuscript, and have added correlations between N2 fixation and primary and bacterial production as requested by another reviewer. Correlations are used in the discussion section and help the interpretation of the results.

18-23 Show 'em or ditch this.
Not sure to get this comment

10 1-5 etc Decimals on figures. It's easy to calculate lots of decimal places on figures, but they clutter up the text. Given the analytical and natural variability of these processes and the analytical processes, how many decimal places are really justified and meaningful?
For N2 fixation rates, given our quantification limit, one decimal only is justified and realistic. This is what is done in the manuscript.

17-18 Realistically, one never really 'measures' a process. Given all of the factors at work, the best we can do is 'estimate' its magnitude. Best to be frank about that.
We agree and replaced 'measurements' by 'estimates'

10-11 16-9 You probably overdid this bit of text. The 'bubble' problem is well known. Best to just say that you used the Montoya method to minimize contamination, but corrected for incomplete dissolution by measuring the 15N/14N ratio. It is interesting that you get higher dissolution in the samples incubated in situ and that needs to be explicitly corrected for
We shortened the text as suggested and removed the following part 'Two methods are routinely used by the scientific community to perform direct N_2 fixation measurements in marine systems: 1) the method developed by (Montoya et al., 1996), which consists of the addition of the $^{15}N_2$ tracer as a bubble in the incubation bottles (hereafter referred to as the 'bubble addition method') and the measurement of the $^{15}N/^{14}N$ ratio of PN before (time zero) and after incubation, 2) the method consisting of adding the $^{15}N_2$ as dissolved in a subset of seawater previously N_2 degassed (Mohr et al., 2010) (hereafter referred to as the '$^{15}N_2$-enriched seawater method'). The second method was developed because the first had been observed to potentially underestimate N_2 fixation rates (Großkopf et al., 2012; Mohr et al., 2010; Wilson et al., 2012) due to the incomplete (and gradually increasing during the incubation period) equilibration of the $^{15}N_2$ in the incubation bottles when injected as a bubble. This results in a lower $^{15}N/^{14}N$ ratio of the N_2 pool available for N_2 fixation (the term A_{N2} used in the Montoya et al. (1996) equation) as compared to the theoretical value calculated based on gas constants, and therefore potentially leads to underestimated rates in some studies (see references above), whereas other studies do not see any significant differences between both methods (Bonnet et al., 2016c; Shiozaki et al., 2015)'.

28 Fig. 1 Bottom: longitude scale is very wrong.
The longitude scale provided here is the conventional one decided for all outpace papers. However we fixed that point.

Would like to see some comparative profiles of measured variables in different regions
The present paper is about the OUTPACE cruise results. We have added N2 fixation ranges from other part of the world for comparison, but this paper is not designed as a review of all the existing literature. Therefore we decided to stay focus on our data.

29 Fig. 2 ODV scales need to be properly annotated.

The x axis label has been added

Fig. 3 Potentially useful, but ... Bottom: should have x-axis scale 0-10 with vertical dotted lines clearly showing natural abundance of 15N (0.367%) and the theoretical 15N/14N ratio if all 15N2 in bubble dissolved. Is the (very) slight mid-water increase in 15N excess statistically valid?
This has been fixed