**Review of Schulz *et al.***

**General comments**

Schulz et al. present a mesocosm study in the Eastern Upwelling System off Peru. The mesocosm approach is relatively novel for studies of this type, although it suffers from its own limitations (e.g., orni-eutrophication). The manuscript is concisely written and proposes that roughly half of the nitrogen that could be exported to depth is denitrified in coastal waters during coastal upwelling. While this study is worthy of publication, I do have some concerns, mostly regarding purging of the samples during the $^{15}N$-labeled incubations, which removes $H_2S$ - an electron donor during chemolithoautotrophic denitrification. Furthermore, since mostly all $O_2$ is removed during purging with He prior to incubating the samples, their measured rates are not representative of *in-situ* conditions ($O_2$ concentrations in bottom waters were generally higher than 20 µmol $L^{-1}$, which is too high for $N_2O$ conversion to $N_2$ (see Dalsgaard et al., 2014)).

**Specific comments**

**Line 9:** "This was in contrast to realized rates in the surrounding Pacific". This sentence is unclear as the author later claim that rates measured as part of the mesocosm experiment and in surrounding waters were comparable.

**Lines 39-43:** This sentence about ocean acidification seems a bit out of place since the connection with OMZs and nitrogen cycling is not clear. I suggest removing.

**Line 73:** How long were the samples stored before analysis?

**Line 74:** $O_2$ concentrations should ideally be monitored during $^{15}N$-labeled incubations using non-invasive $O_2$ measurement technology, such as Oxysense (http://www.oxysense.com/), Pyroscience (http://www.pyroscience.com/) or custom-made sensors (e.g., see Larsen *et al.*, 2016). This is also to ensure that no $O_2$ is infiltrating during the incubations from the stoppers.

**Line 75:** I think such a high $O_2$ offset is problematic and not representative of anoxic coastal waters off Peru – where $O_2$ concentrations are generally well below 10 µM. I suspect that $O_2$ is also introduced during sampling from a Niskin bottle (de Brabandere et al.. 2012). Furthermore, $O_2$ in the nanomolar range has been shown to influence $N_2$ production rates (Dalsgaard et al., 2014). Since their samples are purged with He, the estimated rates are only putative and likely not representative of real *in-situ* conditions. I understand the limitation, but *in-situ* incubations would be preferable.

**Lines 80-82:** Why is this treatment referred to as a moderate treatment if the N:P values between this and the extreme treatment are similar? I would rename this treatment as it is clearly not representative of moderate N loss conditions.

**Lines 82-83:** How did they prevent gas exchange and minimize $O_2$ contamination in these waters during collection/injection? As this is an important detail for their experiment, a brief explanation should be added here (without having to refer to the Bach et al. (2020) paper).

**Lines 87-88:** Why not use $^{15}N$-labeled $NO_3^-$ to measure denitrification rates? Nitrate concentrations are generally higher than $NO_2^-$ and $NH_4^+$ (thus, the substrate is less limiting).

**Line 89:** Why 3 µmol L-1? It seems to be a bit arbitrary.

**Lines 86-93:** On a cautionary note, other studies (de Brabandere et al., 2013; Chang et al., 2014), observed that more $^{29}N_2$ is sometimes produced than could be accounted for assuming a binomial distribution after taking the production due to anammox. They propose that "nitrite shunting" where $NO_3^-$ is converted to $N_2$ completely intracellularly without exchange with the external ambient $NO_2^-$ pool could lead to that $^{29}N_2$ excess. I am curious to know if such $^{29}N_2$ excess was also observed in this study. Using $NO_3^-$ as a tracer could help to account for this process.

**Lines 93-101:** I understand that purging with He is necessary since these are anoxic incubations, but since $H_2S$ was present in bottom waters, removing all gases (including $H_2S$) would underestimate chemoautotrophic denitrification. I strongly recommend complementation of stripped gases involved in N cycling metabolisms.

**Line 99:** What is the recommended flow rate? I assume purging time is adjusted to leave enough background $N_2$ for GC-IRMS measurements?

**Lines 102-103:** At what temperatures were samples equilibrated?

**Lines 110-114:** This approach could potentially affect their rate calculations if "nitrite shunting" produces $^{29}N_2$ excess. See above comment (lines 86-93).

**Lines 113-115:** In figure A1, it seems like there might be an exponential increase from time point 6 hrs and 20 hrs. In this case, I would only use the linear portion for rate calculations. Is this observed for all other incubations? It would have been useful to obtain another time point somewhere in between (at 12 hrs) to better disregard this possibility.

**Lines 134-138:** A more correct approach would be to construct Michaelis-Menten curves and calculate the half saturation constants and maximum denitrification rates from the measured *in-situ* rates (see Michiels et al., 2019).

**Lines 146-147:** Change to: "24 hrs per day x 38 days x 2 (conversion between $N_2$ to N) divided by 3 (contribution of bottom layer water to overall mesocosm volume), ..."

**Lines 189-191:** The oxygen concentrations shown in Fig. 3E are generally above 20 µmol $L^{-1}$, which would be too high for $N_2O$ conversion to $N_2$ (see Dalsgaard et al., 2014 and Frey et al., 2020). Therefore, their measured rates are potential and denitrification was likely only observed because the samples were purged with $O_2$ - removing mostly all $O_2$.

**Line 194:** I don't think it make sense to call this a "moderate" treatment (see above comment lines 80-82).

**Lines 199-200:** Again, these relatively high $H_2S$ concentrations indicate that chemoautotrophic denitrification might be an important process that was not measured (since samples were purged with He before the incubations).

**Lines 211-214:** The near perfect agreement between the two approaches is a bit surprising considering that measured rates are potential and likely not representative of *in-situ* conditions.

**Lines 221-223:** Were $N_2$ fixation rates measured at the same depths than for the denitrification/anammox incubations in the mesocosms?

**Line 266:** The calculated overall nitrogen loss could also be overestimated since *in-situ* denitrification rates were likely lower. Samples collected in the mesocosms and surrounding Pacific waters were purged before the incubations, removing mostly $O_2$ and thus creating conditions more conducive to $N_2$ loss. The $O_2$ concentrations observed in bottom waters were too high for $N_2O$ conversion to $N_2$ (see Dalsgaard et al., 2014).

**Lines 313:** Why did C/N values not increase in that one mesocosm?

**Lines 319-321:** It is also possible that the measured DON pool was mostly recalcitrant, with fast cycling of labile DON.

**Lines 340-341:** Denitrification/anammox linked to microenvironments around particles would not be captured by $^{15}N$-labeled incubations, especially if these are not performed *in situ*.

**Lines 343-345:** It is unclear how $H_2S$ would inhibit anammox in their incubations, since samples were purged (hence $H_2S$ was removed).

**Lines 350-353:** It would be relevant to include these data (i.e., anammox functional marker gene *hzo*) in the manuscript.

**Line 375:** Why is the contribution from the Arabian Sea, where significant N-loss occurs, is not taken into account here?

**Lines 376-379:** I don't think there is anything new in this statement. Due to the large uncertainties associated with these estimates, it is still unclear if the majority of the N-loss occurs in the water-column or sediments.

**Lines 379-380:** Why is export production projected to decrease if upwelling intensity and frequency (and thus nutrient supply) is expected to increase (Hauri et al., 2013 and Wang et al., 2015 papers cited in the introduction)?

**Table 2:** I would rename the "moderate" treatment to "extreme" since the degree of N-loss is similar in both treatments.

It is odd to express individual N-budgets for each mesocosms as negative values and present the mean as a positive value. I suggest renaming these columns N-loss from $^{15}$N-labeled incubations and N-loss from N-budget.

**Figure 2:** What was bottom depth at the mooring site?

**Figure 1A:** It is difficult to tell if the last time point around t = 20 hrs represents an exponential increase (as often observed for $^{15}$N-labeled incubations).

**Technical corrections**

**Line 49:** define nm (i.e., nautical miles)

**Line 63:** change acoording to "according"

**Additional references:**

Dalsgaard, T., Stewart, F. J., Thamdrup, B., De Brabandere, L., Revsbech, N. P., Ulloa, O., ... & DeLong, E. F. (2014). Oxygen at nanomolar levels reversibly suppresses process rates and gene expression in anammox and denitrification in the oxygen minimum zone off northern Chile. *MBio*, *5*(6).

De Brabandere, L., Thamdrup, B., Revsbech, N. P., & Foadi, R. (2012). A critical assessment of the occurrence and extend of oxygen contamination during anaerobic incubations utilizing commercially available vials. *Journal of microbiological methods*, *88*(1), 147-154.

Larsen, M., Lehner, P., Borisov, S. M., Klimant, I., Fischer, J. P., Stewart, F. J., Canfield, D. E., & Glud, R. N. (2016). In situ quantification of ultra-low O2 concentrations in oxygen minimum zones: Application of novel optodes. *Limnology and Oceanography: Methods*, *14*(12), 784-800.