In this manuscript, the authors present newly estimated global ocean N_2O flux to the atmosphere and its confidence interval using observations and two submodels of N_2O production. The paper provides interesting insights but the writing could be improved to make the manuscript clearer. The main problem of the paper, as I see it, is that there are not enough details to assess the validity of the model and results.

We thank the reviewer for the comments. We have tried to clarify the methodology throughout the manuscript.

Below are some major comments and questions, followed by minor edits.

Major comments/questions:

It is unclear how the authors calculate the best estimate of N_2O production using observations (l. 82). How is the range obtained in this case? I thought that the authors might be using the maximums and the minimums of each factor to calculate the range but that does not seem likely.

Errors were calculated with standard error propagation; we added the line: "Here and in the rest of the paper, errors were propagated in the usual way:

error = $(((\text{error of A})/A)^2 + ((\text{error of B})/B)^2 + \ldots)^{0.5} \times A \times B \times \ldots$"

I am having hard time understanding the equation 1. How is this equation derived and why are such large significant figures used? This equation does not account for the latitudinal dependence of pN_2O - wouldn't that be a problem? Isn't it better to use atmospheric model results validated by atmospheric measurements of N_2O?

Eq. 1 is derived from the data in Freing et al. 2009. However, the numbers stated in that paper as the fit to their data are in error, so we here provide the correct numbers as provided by Alina Freing in a pers. comm.. We initially used the numbers exactly as given to us by Alina Freing, but it's true that the number of significant digits is larger than is warranted and we've reduced the significant digits to 7 or 8, so that pN2O is accurate to 2 decimal places.

We added monthly atmospheric measurements at 12 latitudes. Because the observations were not accurate enough prior to 2000 to show a consistent latitudinal gradient and seasonal cycle, the gradient and seasonal cycle were calculated from the data from 2000-2016 and then added to the older global average observations. We added this description to Section 2.6:

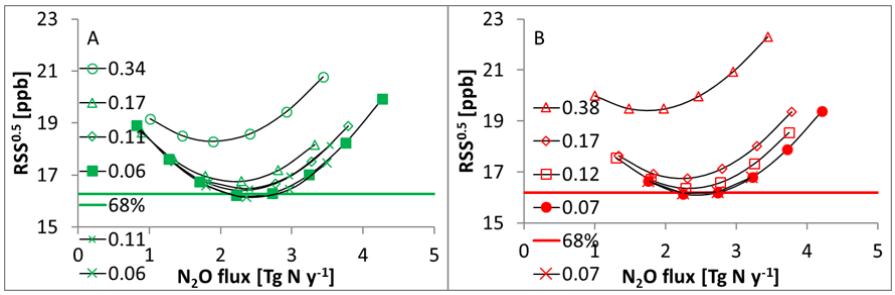
"we also ran a series of simulations with the NOAA pN2Oatm observational data that included seasonal and latitudinal variations. Between 2000 and 2014, we used the monthly observations for the 12 available latitudes. Monthly anomalies relative to the global average were calculated at each available latitude from the 2000-2016 observations. These were added to Eq. 1 from 1965 and 1976, and to the global average observations between 1977 and 1999. In the model simulation, the data were linearly interpolated between the 12 latitudes and monthly observations."

And we added to Section 3.2:

"When we used observed atmospheric pN2O that varied with latitude and month (see Section 2.2) the result was essentially the same, with an N2O flux of 2.4 ± 0.3 Tg N y-1 for the diagnostic sub-model and 2.6 ± 0.3 Tg N y-1 for the prognostic sub-model (data not shown)."

Although not included in the manuscript, we here include a modified Fig. 11 (was Fig. 10 in the submitted manuscript), with the additional simulations using the NOAA pN2Oatm observations shown as crosses (at two low O2 production rates for the diagnostic model and

at the optimum net low O2 production rate for the prognostic model), which shows that when we used the observed atmospheric pN2O the results were essentially the same:



See also the reply to the comment by reviewer 2 on Line 219-220.

I think there might be a mistake in equation 2. Otherwise, I do not see how a value of 2 could mean that the model deviates form the observations by a factor of 2 in either direction. 10^(10log2) = 1024 and it is nothing close to a value of 2. Please explain.

Perhaps, the standard mathematical notation (summation and the number of observations n rather than "average") would be more appropriate here.

The 10 before log indicated that it's the 10-base logarithm. This has been corrected to $\log_{10}$, $\log_{10}(2)=0.31$ and $10^{\wedge 0.31}=2$. We've converted the manuscript to Latex, which allows a subscript inside a superscript, which makes this distinction more clear. We've changed the formula to Σ…/n.

It would be useful if the N_2O flux calculation in section 2.7 is explained in a little more detail, rather than stating that it "is calculated with the piston velocity from Sweeney et al. (2007)." I am not familiar with this calculation and would love more explanations on how the ocean N_2O flux is estimated but the Sweeney et al. (2007) is not listed in the references either.

We've added the equation for the N2O flux calculation, including the piston velocity and the reference to Sweeney:

"N2O flux (=air-sea gas exchange) is calculated as:
N2O flux = (pN2Oatm*K0*(1-p_watervapor)-pN2O)*piston_velocity*{660/Schmidt_number_N2O}$^{0.5}$*(1-ice_cover)
, in which K0 is the solubility {WeissPrice80}, p_watervapor is the water vapor pressure {Sarmiento92}, piston velocity = 0.27*(wind speed)^{2} {Sweeney07}, which is optimised for use with the NCEP reanalysis data used here, the Schmidt number for N2O was taken from {Wanninkhof92}, and the ice cover is calculated by the sea ice model LIM2."

I am not sure how equation 3 is used to determine the global air-sea flux of N_2O that best fits the ΔpN_2O data, if RSS/RSS_{min} just depends on the number of observations. I do not understand how different model simulations would have different values of RSS/RSS_{min} if the number of observations is the same.

It is not n that varies but rather RSS varies as the results of different model simulations are compared to the same observations. We added information about regridding and the calculation procedure in Section 2.3, Eq. 4 in Section 2.8, and added clarifications to the legend of Fig. 9:

"The 1σ confidence interval, where RSS equals the value calculated from Eq. 3, is indicated by the horizontal lines. A) diagnostic submodel, each point represents a simulation with a different low O2 slope, B) prognostic model, "no c" is with no N2O consumption i.e. net

production = gross production. All other lines have a constant gross production, and net production varies with different N2O consumption rates. Range of parameter values is given in Section 8.7 of the supplementary material." and of Fig. 11: "MSE^0.5 for the two N2O submodels compared to the ΔpN2O database as a function of global N2O flux at different (net) N2O production rates in the low O2 regions. A) diagnostic submodel, the four lines represent the four best low O2 production rates from Fig. 9A, each point represents a simulation, different symbols indicate different low O2 slopes, points with the same symbols have different oxic N2O production slopes. B) prognostic submodel, the four lines represent the optimised net production rates at the four best gross production rates from Fig 9B, points with the same symbols have different N2O slopes for nitrification."

As for equation 4, I think that its application should be described within the methodology section, rather than just mentioning a little in the discussion section.

Since the F-test at large sample size is insensitive to non-normal distributions we have deleted the equation and accompanying text.

Also, how did the authors optimize various model parameters? And is it not a problem that the optimized oxic $\Delta N_2O/AOU$ slope of 12.7 μmol N_2O (mol O_2)$^{-1}$ is so different from the global average given earlier in lines 77-78 (81.5 ± 1.4 nmol/mmol)? What is the value for this parameter in the prognostic model?

The observed slope of 81.5 μmol N_2O (mol O_2)$^{-1}$ in figure 3 is a weighted average of the low O_2 slope and the oxic slope. The optimised slopes in the model are 1700 μmol N_2O (mol O_2)$^{-1}$ under low O_2 and 12.7 μmol N_2O (mol O_2)$^{-1}$ under oxic conditions. For the weighted average of these model slopes to equal the observed slope of 81.5, the fraction of N_2O that is produced by the low O_2 regions would need to be 4.1% (=(81.5-12.7)/(1700-12.7)). This is close to the 6% for the diagnostic model and 4% for the prognostic model that we find. Since this 4.1% is simply a sanity check that the optimised model does a reasonable job of reproducing the data, but is not an independent estimate, we have not added this calculation to the paper.

The slopes for the prognostic model are given relative to the substrate for each pathway (NH_4 for nitrification, NO_3 for denitrification). To allow for an approximate conversion to O_2 specific slopes (i.e. under the simplifying assumption that NH_4 and NO_3 are consumed at the same place where they are produced), we've added to section 2.5 that: "Phytoplankton (and all other organic matter) have a fixed C:N:O_2 ratio of 122:16:-172." From this it can be calculated that the prognostic model oxic slope of 123 μmol N_2O (mol NH_4^+)$^{-1}$ approximately converts to 11.5 μmol N_2O (mol O_2)$^{-1}$. Because reviewer 3 was not entirely clear whether denitrification in the model is actual denitrification using NO3, we did not add this O2 based slope in the manuscript, as it would add to the potential confusion.

Minor comments

1. L. 24 "It also currently" a • "It is also currently"

Changed.

2. There are several places in the text, where more detailed or clearer explanations would help readers understand the paper better. For example, l. 53-56 is unclear what the sentence means. Do the authors mean that $\Delta N_2O/AOU$ slope becomes negative under suboxic conditions and that leads to the ambiguity of how much N_2O is produced in this region? Please clarify.

We've expanded on the ambiguity to clarify potential reasons for it.

3. L. 71 "observationally derived" a • "observationally-derived"

The Chicago manual of style says not to hyphenate adverbs ending in -ly.

4. L. 75 Since not all readers of this paper are experts in oceanic biogeochemistry, it would be helpful to explain that the f-ratio is the fraction of total primary production by nitrate.

Added.

5. L. 79 What is the "-O_2:C ratio"? What is the dash for?

We've added that "(the - sign indicates that O2 is consumed as CO2 is produced)"

6. L. 233 "N cycle based" a • "N cycle-based"

Changed to N-cycle-based.

7. L. 242-246 "This estimate…" run-on sentence and needs to be rewritten.

This was split into two sentences.

8. L. 263-267 "It should also…" run-on sentence and needs to be rewritten.

This was rewritten and split into 4 sentences.

9. L. 286 "140 pm" a • "140 ppm"

Changed

10. L. 290-294 "On the one hand…" run-on sentence and needs to be rewritten.

This was split into two sentences.

11. N-cycle data database used in this paper are shown as embargoed in the data source pointed by the authors (https://www.uea.ac.uk/green-ocean/data). Will the data be publicly available?

The data have now been made publicly available.