We thank the reviewers for their thoughtful and constructive comments which will greatly help us in preparing a revised manuscript. Here we provide our replies (in bold text) to their comments (in plain text).

Reviewer #1

The manuscript presents a new estimate of oceanic denitrification in the water column and in the sediments from a coupled 3D ocean circulation-biogeochemistry model that is optimized against physical and biogeochemical observations, and argues for a balanced oceanic nitrogen cycle. Denitrification obviously plays a major role in the cycling of nutrients and carbon in the ocean. However, estimates of the rates of oceanic denitrification both in the water column and in the sediments are plagued by large uncertainties - to the point that it is not even clear whether the oceanic nitrogen budged is balanced (for example by internal biogeochemical feedbacks) or not.

Only recently, estimates obtained with different methods have started converging to a narrower range of values. The uncertainty is particularly acute for benthic denitrification, which, by means of simple considerations on the isotopic composition of oceanic nitrate, should be between 1-4 times larger than water column denitrification. In the literature, sediment denitrification estimates vary 2-3 folds among each other. Such estimates rely on either poorly-constrained benthic models, or on the combination of water column denitrification rates with global isotope data, an equally uncertain method. In this perspective, the new estimate by DeVries and coauthors, which elegantly integrates informations on nitrogen deficit and isotope distributions, is an important contribution that brings additional evidence for a balanced nitrogen cycle.

I found the method used by DeVries and coauthors quite sound, as it integrates our knowledge on ocean circulation dynamics (momentum and mass conservation) with observed tracer distribution (temperature, salinity, nutrients and N isotopes) in an optimization framework. Major model limitations could include: (1) the use of linearized momentum equations, and the assumption of a steady state circulation and tracer distributions; (2) data availability; (3) a simplified representation of biogeochemistry; (4) computational constraints. I am not an expert in data assimilation techniques, but the circulation model used by the Authors (although in different incarnations) has been thoroughly described and validated in previous papers. Similarly, I cannot comment much on the assumption of linearity in the momentum equations, except that it seems acceptable for the type of large-scale processes addressed here (and probably not worse than the common use of coarse-resolution general circulation models for biogeochemical studies). The assumption of a climatological steady state appears adequate for the scope of the study. The biogeochemical model adopted is rather simple, but analogue to widely tested biogeochemical formulations (e.g. OCMIP-type "restoring" models). Obviously, the simplicity of the model allows for a minimal number of parameters that is functional to the optimization approach. That said, I must note that fundamental processes such as non-Redfield organic matter production and nitrogen fixation are explicitly considered by the Authors. The inverse-model approach does not grant that the optimization does not end up compensating for model deficiencies. As an example of this kind of compensations, one could ask whether a reduced shallow denitrification due to unresolved shelves might force the optimization to increase benthic denitrification in the rest of the ocean in order to achieve N* and N-isotope consistency with the data. However, this sort of problem is common to any modelling effort. Finally, data availability and computational constraints might be at the base of some of the technical choices of the study, e.g. the 2-step optimization of the circulation/biogeochemistry first and of the nitrogen cycle second, and the subset of parameters chosen for the optimization. More isotope measurements would certainly not hurt, but the major locations where water column denitrification is known to happen, as well as regions representative of larger areas of the ocean basins, are adequately covered in the study.

The manuscript is well organized and written, and the assumptions, the model limitations, and the sources of uncertainty are generally clearly addressed. I consider this study an important and useful contribution - both in terms of the results and of the methodology - and I recommend its publication in Biogeosciences. Overall, I think that a few minor points could benefit from clarifications:

(i) Model formulation. What is the value of the depth attenuation coefficient for particles (b) used in the biogeochemical model? This is perhaps the most important parameter for remineralization as it controls the proportion of particulate organic matter that is remineralized in the water column and the fraction that reaches the bottom. As such, the choice of this parameter has a first impact on denitrification by controlling the particle flux that fuels sedimentary denitrification. Furthermore, we know that the attenuation of the particle flux within suboxic regions is substantially lower than in the open ocean (approximately 0.3-0.4 vs. 0.8-1.0, e.g. Martin et al., 1987; Devol and

Hartnett, 2001, Van Mooy et al., 2002). This implies that less export is remineralized inside suboxic waters. In fact, a recent study by Bianchi et al. (2012) showed that both water column and sediment denitrification are quite sensitive to the value of b inside and outside suboxic waters. Thus, I'm a little surprised that (1) a uniform value of b is used for oxygenated and suboxic waters; and (2) b is not included in the set of parameters optimized against N* and N-isotope measurements. I suspect that the lack of oxygen-dependence for b, and the lack of optimization against nitrogen tracers, might be due to data limitation (do nutrient measurements allow to solve for an oxygen-dependent b, when so many other factors control particle remineralization?), as well as limitations in the optimization procedure (that is, first using PO4 to solve for the optimal circulation, second solve for the optimal nitrogen-cycle model parameters) that require the same b value for particulate organic phosphorous and nitrogen flux attenuation (otherwise the N:P stoichiometry of remineralization would vary with depth). However, given that b controls the model partitioning of remineralization between water column and sediments, I feel that a discussion of the choices of keeping b constant, and not including it as part of the nitrogen-cycle optimization, is needed.

The value of b used in the model is 0.77 (for the model in which σ_{DOP}=2/3) or 0.79 (for the model in which σ_{DOP}=1/2). This value is based on the results from a joint inversion for the circulation parameters of the model and the coefficient b against global data sets of temperature, salinity, radiocarbon, and PO_4 (i.e. step 1 of the inversion).

We use a globally uniform value of b primarily for simplicity, and because we were able to achieve a good fit to observed PO_4 using a spatially uniform b. We assumed that particulate organic nitrogen remineralizes following the same depth profile as that for phosphorus, and so we use this same value of b in the optimization of the parameters of the N cycle model (i.e. step 2 of the inversion). If we were to allow it to vary as part of the set of parameters optimized in the N cycle model, it would likely revert back to the value found in step 1 of the inversion, since at a global scale the remineralization of organic N and P is likely to occur at similar rates.

By performing the inversion in two steps we are implicitly assuming that the PO_4 observations tell us all we need to

know about the remineralization rate of organic matter, and that the N* data do not contain any information about circulation rates. This is obviously a simplification, but one that we had to make in this case in order to make the problem computationally tractable. In the future we hope to be able to perform the inversion in one step with a spatially variable remineralization rate for both organic phosphorus and organic nitrogen.

It is hard to judge exactly what the effect of assuming a uniform value of b is on our solution. The value of $0.77-0.79$ used is in line with observed values of this parameter outside of suboxic zones. Since most benthic denitrification occurs outside the suboxic zones (which are very small), ignoring the spatial variability in b is unlikely to have a large effect on the modeled rates or spatial pattern of benthic denitrification. On the other hand, it could have a significant effect on the modeled denitrification rates in the suboxic zones. There is some indication of misfits between modeled and observed N* within the suboxic zones (see Figure A1). For example, the modeled N* peak is too deep in the Arabian Sea, while the very low near-surface N* values in the ETSP are not captured by the model. This misfit could be due to mis-specification of b, but it could also be due to inadequate circulation (as indicated in the CFC-11 depth plots). For example, the inability of the model to match the near-surface negative N* values in the ETSP could be due to inadequate upwelling and too-low productivity, which would also explain the too-high CFC-11 concentrations in that region (Figure A1).

We should not that the globally integrated water-column denitrification rates inferred here are consistent with those inferred by DeVries et al. (2012). That study did constrain the circulation in the OMZs using CFC-11 observations, and did take into account uncertainty in the production and remineralization rates (specifically the value of the Martin coefficient b) within the suboxic zones. Although the value of b differed from that used here based on global PO_4 fits, the total rates were not significantly different.

(ii) Figure 3.b: the model overestimates the magnitude of N* in the thermocline and deep ocean - with a bias of up to -1 umol. This misfit is described in section 2.3. Whereas the overall denitrification signal is about 5 times larger, the implications of this bias could be further discussed. For example, does it imply that the model overestimates the strength of denitrification in the thermocline and in the abyssal ocean by up to 20 % at those depths over the whole ocean? This seems at odds with the low denitrification rates found in the water column and in particular in the sediments (relative to existing estimates).

As the reviewer correctly points out, the model-data misfit in N* points to slightly too much denitrification in the thermocline and deep ocean. We think it likely that there is a deep bias in sedimentary denitrification in the model. We believe that this is primarily due to coarse resolution of the model - which cannot resolve all of the continental shelf areas, and does not allow high enough rates of coastal production to be simulated. So it is likely that denitrification in shallow shelf areas is too low in the model, and denitrification in the slope and deep sea areas is too high. If denitrification in shallow areas produces less of an imprint on the mean ocean $\delta^{15}\mathbf{NO}_3$, this could produce a slightly low bias in the total benthic denitrification rate found by the model.

(iii) I found the discussion in section 4.1 useful, but rather technical. I suggest that the Authors introduce the paragraphs addressing the effects of denitrification on N* and N:P ratios with a couple of sentences that briefly introduce why these effects are important to a broader audience.

Thank you for pointing this out. We will address this in the revised manuscript.

(iv) Sedimentary denitrification model. While I have no reason to prefer a model formulation over another, given the empirical nature of the sedimentary model itself, and its subsequent optimization against data, the authors could be more generous in describing the assumptions behind it. In particular, the only reference given is the study by Middleburg et al., 1996, which provides quite different empirical relationships. For example Middleburg's relationships show a non-linear dependence between denitrification and particle fluxes, while DeVries and coauthor assume linearity (equation A11). Also, why a hyperbolic dependence is used for NO3, and a hyperbolic tangent for O2 (equations A13-14)? Again, these are technical choices but more discussion could be useful.

We use a linear formulation for benthic denitrification because highly non-linear expressions introduce challenges in applying Newton's method to solve the steady-state model equations, and in applying the quasi-Newton method to find the optimal model parameters. However, both the linear fomulation in this paper and the non-linear formulation of Middleburg et al (1996) produce similar benthic denitrification rates and depth pattern (T. Weber, personal communication).

We used a hyperbolic tangent function for the O_2 dependence of benthic denitrification because it was unclear to us how the transition to low-O2 regions should be modeled (i.e. should it be an abrupt transition or a smooth transition). Ultimately, we found that the parameter controlling this (K_{O_2}) is poorly constrained (see Table B1), indicating that we probably could have used a hyperbolic dependence for both O_2 and NO_3 dependence of benthic denitrification. However, this choice is not likely to influence the model results.

We should also point out a typo in Equation A13. The 1 should be outside the parentheses (not an argument of the tanh function). We will correct this in the revised paper.

Technical comments:

(i) page 14015, line 21. The delta notation should express the ratio between 15NO3/14NO3 and a standard (e.g. atmospheric N2).

Thank you for pointing out this error.

(ii) equation A6. What is Tmax, and what value is used?

Tmax is the maximum modeled surface temperature (about 31°C), which is included to ensure that the maximum of the exponential expression is 1.

Reviewer #2 (N. Gruber)

Summary

DeVries and coauthors use an inverse approach that combines observations of various tracers and an ocean circulation model to estimate the rate of marine denitrification. They find a global denitrification rate of between 120 to 240 Tg N yr^{-1}, which is at the lower end of all estimates, but in line with the most recent ones. About one third of the total rate is driven by water column denitrification, and the other two-thirds by benthic denitrification. The implied ratio of benthic to water-column denitrification of about 2 is also lower than the original estimates, and implies a relatively modest global efficiency by which the isotopic signature of water column denitrification imprints itself on the global NO_3 pool.

Evaluation

Each year denitrification removes several hundred Tg N yr^{-1} from the fixed nitrogen pool in the ocean, stripping the ocean from this essential nutrient. Yet the views on how large this sink actually is still diverge substantially, although most recent estimates tended to cluster on the lower end of the spectrum.

DeVries et al. now add a very powerful and important new estimate to this discussion. They use, for the first time, a global 3-D model, and assimilate a suite of N-related tracers to estimate separately the rates of water column and benthic denitrification. They thereby demonstrate the importance of the previously identified dilution effect in determining the global ratio between water column and benthic denitrification. I particularly like the insightful discussion of the factors that control this dilution effect and hence the benthic to water column denitrification ratio, and also the discussion of how denitrification impacts the determination of the N:P remineralization rates from observed NO_3 and PO_4 data.

The study was carefully and insightfully designed, the paper is well written and illustrated, and the results are clearly novel, interesting and important. This manuscript is therefore very well suited for publication in Biogeosciences. I have many comments, but none of them is of a fundamental nature. They are rather intended to make an already excellent study (hopefully) better.

I list here the major comments, while I discuss the minor (general and specific) ones below. All of the major comments deal with various error sources whose potential contribution to uncertainty could be better discussed. At the moment, the uncertainty section considers essentially just the "internal" uncertainties, i.e., those emerging from the assimilation system, and pays limited attention to the "external" errors, i.e., biases, particularly those that are of structural origin.

(i) Circulation model error: The results critically hinge on the models ability to correctly capture the circulation of the oxygen minimum zones, as it is the relative consumption of nitrate in these regions as well as the "efficiency" by which this signal is mixed out into the rest of the ocean that is critical for determining the dilution effect. I therefore consider it important to learn more about how well this model is able to capture the circulation of these regions. Global coarse resolution model have notorious problems in these regions, and it is not clear that a data constrained model will necessarily do better. In this regard, I was a bit surprised to read that the authors used here a version of the circulation model that was not optimized with CFCs, but "only" with T, S, and radiocarbon. While the latter is certainly a very good constraint for the deep ocean, it is not that well suited for constraining thermocline rates. There is some indication of the thermocline circulation potentially being a problem in that the optimized profiles of N* differ substantially from the observed one in the thermocline across the Indo-Pacific. It thus seems to me that this aspect deserves a deeper discussion.

In this study we did not optimize the circulation with CFCs since we thought it more important to optimize to phosphate. It is necessary to obtain a good fit to PO_4 in order to interpret the N* data, which represent fairly small deviations from a linear relationship with PO_4. Because of the additional computational cost associated with assimilating PO_4, we were unable to also assimilate the CFC observations. In the future, we plan to obtain a circulation optimized with both CFC and PO_4 observations.

When compared to observed CFC-11 concentrations, the model does show some deficiencies in ventilating the suboxic zones. In particular, the Arabian Sea OMZ is apparently too weakly ventilated near the surface and too well ventilated at depth, while the ETSP is too well ventilated throughout (see Figure A1). This could cause some of the errors in the model-data N* misfit in these same suboxic zones (see Figure A1),

although errors in the biogeochemical parameters (e.g. the depth attenuation of organic matter remineralization) cannot be ruled out (see response to Review 1 above). In particular, we think it likely that modeled upwelling in the ETSP is not strong enough, due to the relatively coarse resolution which is inadequate to fully resolve coastal upwelling. This would cause both the too-high CFC concentrations (near-surface waters are too young) and the too-high N* concentrations (sub-surface waters with large N deficit are not upwelled strongly enough into near-surface waters, reducing productivity and also reducing denitrification).

Irrespective of these circulation errors, there are two important points to be made. First, the water-column denitrification rates derived here (50-77 TgN/yr) are very similar to those derived by DeVries et al (2012) (66+-6 TgN/yr) using a model that was constrained by CFCs and that used N2/Ar data to constrain denitrification rates. The agreement between the two independent estimates indicates that the range of water-column denitrification rates derived here is appropriate. Second, the overall good fit to the suboxic zone N* in terms of fractional consumption (Figure 2d) indicates that the magnitude of the dilution effect should be approximately correct.

(ii) Data error: The other main ingredient of any data assimilation system are the data. My understanding is that the authors are using the objectively mapped N* data of the World Ocean Atlas. These data underestimate the extent and magnitude of the low N* in the oxygen minimum zones (see e.g., Eugster and Gruber, (2012)), likely due to the strong smoothing that was applied when this data product was produced. Presuming that a substantial fraction of the results are driven by the model trying to match the low N* data in the oxygen minimum zones, any errors in the data have a direct effect on the results. I also wonder why the N2/Ar data used in deVries et al. (2012) were not included here as an additional constraint.

It is true the objectively mapped N* data do smooth out the very low N* concentrations in the OMZs. Unfortunately, it is difficult to address this issue in the context of the model because even if the raw observations are binned to the model

grid, the relatively coarse resolution of the model necessitates that a large degree of smoothing will take place. One could hypothesize that smoothing out the very low N* data will cause the model to underestimate water-column denitrification rates, but a stronger N consumption in suboxic zones would also cause a stronger dilution effect leading to lower benthic denitrification rates (to match the mean ocean δ^{15}NO$_3$). It is not clear what the ultimate effect would be on the total denitrification rate.

This issue of smoothing, along with the circulation issues (above) and the resolution of shelf areas (below), is a very good reason to increase the resolution of the model in future studies. We are currently working on this.

(iii) Shallow seas: With much of benthic denitrification occurring in shallow seas that are poorly represented in the relatively coarse resolution model, it is unclear how this structural error imprints itself onto the final results. Our experience using a structurally much simpler model, but essentially the same data constraints (Eugster and Gruber, 2012) leads me to believe that this may turn out to be rather unimportant source of error, as the global results are strongly driven by two numbers, i.e., the water column denitrification rate, and the global mean 15N. But it would be useful to know more about this than the somewhat "ad hoc" argument that the likely underestimate of benthic denitrification may be compensated by the lack of consideration of the riverine input of N.

We agree that the large-scale balance between benthic and water-column denitrification is primarily constrained by the water-column denitrification rate (specifically the fractional N consumption in suboxic zones) and the mean ocean δ^{15}NO$_3$. However, the poor representation of shelf areas in the model, as well as the poor representation of coastal productivity, probably causes the model to underestimate denitrification in shelf areas, which it compensates for by increasing benthic denitrification in the deep ocean. That there is too much deep benthic denitrification can be seen in the depth profiles of N* in the Indo-Pacific basin, which shows too low N* in the deep ocean (see response to Reviewer #1 above). It is not totally clear how shifting the depth structure of benthic

denitrification would affect the mean ocean $\delta^{15}NO_3$ in the 3-d model. Increasing the resolution of the model in the future will help us to answer this question.

(iv) Atmospheric deposition and Riverine input: These are two important sources of fixed N to the ocean, perhaps as large in magnitude as water column denitrification. In addition, these sources might have undergone a substantial change over the anthropocene, with some studies suggesting a doubling of the overall input. This raises two questions: First, in what way will the lack of consideration of these two fluxes impact the results? Second, how will the large transient in these fluxes interfere with the essentially steady-state assumption that underlies this inverse modeling system?

We agree that the sources of N from atmospheric deposition and riverine inputs are both likely to be important to the global N budget. They may also have a detectable effect on the distributions of N* and $\delta^{15}N$, especially in the surface ocean where these fluxes occur. Our neglect of these signatures is largely based on the fact that our focus is not on surface tracer distributions, but rather on their subsurface modification during denitrification. In the case of N*, whatever spatial pattern these fluxes impart to surface waters is achieved by restoring toward observed NO_3 and PO_4 independently. Their effect on the N reservoir will be implicitly included in our "N fixation" term, which serves primarily to close the N budget while doing little to modify subsurface tracer distributions. In the case of isotopic constraints, the effect of these surface fluxes is not similarly accounted for, since an observed distribution of $\delta^{15}N$ is not enforced through restoring, as it is with N*. Given the overwhelming effect of fractionation during NO_3 assimilation on the surface $\delta^{15}N$ distribution, it seems unlikely to us that the spatial pattern of $\delta^{15}N$ from these exogenous sources would be very important. Instead, their more important effect is on the global mean NO_3 $\delta^{15}N$. If the isotopic signature of those fluxes is on balance not significantly different from that of N fixation (e.g. Brandes and Devol, 2002) then these inputs could again be considered part of the "N fixation" term that closes the budget. Given the uncertainty surrounding the isotopic ratio of atmospherically deposited N, this assumption must be taken

as provisional. We will discuss these caveats more fully in the revised manuscript.

The question of transients in these fluxes is more difficult to address. If the oceanic N* data is "contaminated" by recent anthropogenic N inputs, then the results may be affected, but it is difficult to know in what sense. We prefer not to speculate in this regard due to the numerous feedbacks on N fixation, denitrification, and production known (or believed) to operate in the ocean and which our model is not adequate to address. Suffice to say that these issues are indeed unresolved and should be addressed in future studies.

Recommendation

I recommend acceptance of this manuscript after a minor to moderate revision. I particularly recommend that the authors extend the discussion of the potential biases in their estimates emanating from structural errors in the assimilation system.

General (minor) comments

Anammox is not mentioned anywhere in the whole manuscript. With some authors arguing that this process represents a large sink for fixed nitrogen in the ocean, it behooves the authors well to discuss this process and what it means for the interpretation of their results.

Thanks for pointing this out. We will add a discussion of anammox in the context of our results in the revised manuscript.

The authors should clarify better the similarities and differences of this study with the recently published deVries et al. (2012) paper in Nature Geoscience. Although this is partially done, it would be helpful for the non-expert reader to be provided with a succinct summary.

This will be clarified in the revised manuscript. The study of DeVries et al (2012) in Nature Geoscience used observations of excess N_2 gas from within the three main open-ocean OMZs to derive a probabilistic estimate of the global marine water-column denitrification rate. They then used published estimates of the ratio of benthic to water-column denitrification to derive a probabilistic estimate of total marine denitrification rates. The denitrification rates derived by DeVries et al (2012) are independent

of the rates derived in the present study (although a similar circulation model was used in both studies), and thus the agreement in water-column denitrification rates in the two studies supports the robustness of these estimates.

I admit that this is self-serving, but the recently published article by Eugster and Gruber (2012) in Global Biogeochemical Cycles addresses many similar issues (e.g., global rates, dilution effect, benthic to water column denitrification ratio) on the basis of a fundamentally similar approach. Thus it would make a lot of sense to discuss these results in the light of the findings presented here. I found it very intriguing that the global rates turn out to be rather similar and also the benthic to water column denitrification ratio is remarkably close. Is this a sign of robustness in these findings, given the very different nature of the underlying circulation models (3D versus box model), or is this just coincidence?

The box model study of Eugster and Gruber (2012) came out after this discussion paper was published and so we were not able to include discussion of it in the present paper. We will do so in the revised paper.

It is interesting that the two studies arrive at similar overall denitrification rate (107-188 TgN/yr in Eugster and Gruber, 120-240 TgN/yr in this study) and benthic to water-column denitrification (1.6-2.0 in Eugster and Gruber, 1.3-2.3 in this study). We think that to a degree this represents the robustness of the results. The ratio of benthic to water-column denitrification (B/W) is determined to first order by the mean ocean $\delta^{15}NO_3$ and the fractional consumption in suboxic zones, and the agreement between the box model and GCM results indicate that details of the ocean circulation are probably a second order influence on B/W. On the other hand, we think that the total rate of denitrification is likely to be more sensitive to details of the ocean circulation, which influences the pattern and magnitude of biological production, the extent of suboxic zones, and the spatial distribution of remineralization. Therefore, the agreement between the box model and GCM in this regard may be partly fortuitous.

Specific (minor) comments

- section 2: inverse nitrogen model. I am wondering how the variations in the N:P uptake ratios in the Southern Ocean are dealt with? I presume that the simultaneous restoring of N and P toward observations takes care of this, and that the particular N-fixation parameterization (linked to atm. Fe deposition) avoids the diagnosis of elevated N-fixation in this region. As Deutsch and Weber investigated this issue in other publications, it would be good to know in a more explicit manner how this is considered. This may be especially important for determining the preformed N* in the thermocline of the Southern hemisphere.

Yes, spatial variations in the N:P ratio are implicit in the nutrient restoring parameterization for production, in which the uptake of NO_3 and PO_4 is computed by restoring to annual mean observations of each respective nutrient (e.g. equations (A3)-(A4)). The parameterization for nitrogen fixation (Equation A(6)) includes the effects of limitation by light, iron, and temperature, and inhibition at high NO_3 levels. Thus N-fixation in the Southern Ocean is very low (total of ~4 TgN/yr of fixation south of 40°S).

- section 2.1: If I am not mistaken, this is an annual mean model. This should be mentioned here explicitly. I don't think that this is extremely critical here, but in the real ocean, seasonal variations in physical supply and nutrient drawdown are correlated, leading to covariances that are not captured by an annual mean model.

  Yes, the model is steady, and does not include a seasonal cycle or interannual variability. This will be explicitly mentioned in the revised text.

- section 2.2: Optimization and appendix B: It would be useful to know how the authors ensure that their optimization method is not falling into a local minimum. With this being a highly non-linear problem with likely a large number of local minima in the cost function, this can easily happen.

  The optimization algorithm used (a quasi-Newton algorithm) cannot distinguish between local and global minima. Due to the large number of control parameters (Table B1) and the computational cost of each model solution, it is not possible to apply a global minimization algorithm to the present model. We did however apply the optimization scheme to 48 different versions of the model, and most parameters show a relatively small posterior uncertainty indicating a strong

minimum (Table B1). Some parameters have a large posterior uncertainty indicating a weak minimum, which might reflect multiple minima, but more likely reflects a broad, weak minimum indicating that these parameters are poorly constrained.

- section 2.2: Cost function: I also think that it would be useful to be more specific and explicit about the formulation of the cost function and add also some details already in the main text. I presume that the authors do not include a regularization (or Bayesian) term (e.g., by penalizing deviations from the initial guess), but I wasn't sure from reading the text.

  We do include a term in the cost function that penalizes deviations from their initial guess. This is necessary to provide some guidance to the optimization routine about the acceptable search space. However, the weight of this term was kept small relative to the weight of the terms measuring the misfit between modeled and observed N* and δ^{15}NO$_3$. See also reply to comment below.

- section 2.3, lines 7-8: "good fit". I would love to see more sensitive measures of model data misfit than a plot of observed vs measured N*. In particular, one wonders about the regional distribution of the residuals. Some of this is shown in Figure 3, but in a highly aggregated manner.

  To address this we have included Figure A1 showing the model-data comparison for N*, δ^{15}NO$_3$, and CFC-11 in the suboxic zones.

- section 2.3, lines 13-14, "relative nitrate consumption". This may not be that relevant in this section, but it will be later in the discussion (p14026, lines 1). So I raise it here already. The equation fc = 1-NO3/(16×PO4) is not really correct for estimating the degree of nitrate consumption in the oxygen minimum zones. This equation works only if preformed N* is 0. Otherwise, the relative degree of nitrate consumption needs to consider the preformed value of NO3 and PO4 explicitly. Based on our own calculations in Eugster and Gruber (2012), we found the preformed N* value to be important for the calculation of the relative N consumption inside the oxygen minimum zones.

Please see reply to comment below.

- section 4.2: I commend the authors for this very useful discussion.

  Thank you.

- section 4.2, p14026, lines 1ff: See above comment on the definition of fc. Despite my concern, the approximation seems to work quite well, but I would love to see this calculation repeated with a more exact definition of the relative nitrate removal. One of the reasons for raising this issue is that we found in Eugster and Gruber (2012) only a moderate relationship between relative nitrate removal and the magnitude of the inversely estimated benthic to water column denitrification rates across the 2500 circulation configurations we considered.

  This point is a bit subtle. The relationship in Figure 6 is indeed an approximation, and the definition of f_c is only one source of uncertainty. If we use the mean ocean ratio of N:P (about 14:1) instead of 16:1, the relationship degrades (predicted B/W values are higher than observed). However, as mentioned in the text, it is difficult to know how to derive a single value for f_c or for the preformed ratio of N:P. Also, the formula does not take into account fractionation due to uptake, another confounding factor.

  The main point we have tried to make with Figure 6 is that the dilution effect reduces the isotopic enrichment effect of water-column denitrification, by a factor that seems to be proportional to the fractional consumption in the center of suboxic zones. The relationship is only an approximation, and clearly more work needs to be done to fully understand how the magnitude of the dilution effect is set. We prefer to address that issue with a higher resolution model with greater fidelity in the suboxic zones.

  In our solutions, we also do not see a very strong relationship between f_c and the ratio of benthic to water-column denitrification, although it is generally a positive correlation. However, the relationship of equation (2) shows a quite good fit.

- Appendix B, lines 19-29. "cost function". See also comment above. In addition, it wasn't entirely clear to me how the different constraints were weighted relative to each other. In addition, the data are highly non-randomly distributed (especially the 15N data), so that some adjustment might have been necessary. Finally, the authors write "This choice is made so that the primary factor controlling the final value of the parameters is constraints provided by the [..] observations". This is unclear to me.

We apologize for the confusion. We did account for the number of observations when creating the cost function. The exact definition of the cost function is

$$\begin{aligned} c &= \frac{1}{N_{N*}\sigma_{N*}^2}\sum(N^*(mod) - N^*(obs))^2 \\ &+\frac{1}{N_{\delta^{15}NO_3}\sigma_{\delta^{15}NO_3}^2}\sum(\delta^{15}NO_3(mod) - \delta^{15}NO_3(obs))^2 \\ &+\frac{1}{N_p\sigma_p^2}\sum(p_{pos} - p_{pri})^2, \end{aligned} \quad (1)$$

where N_{N*} ($N_{\delta^{15}NO_3}$) is the number of grid cells with N* (δ^{15}NO$_3$) observations, and N_p is the number of control parameters (see Table B1). p_{pos} represents the (posterior) model parameters, and p_{pri} their prior values, which were used as the initial guess (see Table B1). We chose σ_p^2 to be very large for all parameters so that the model solution was not biased toward our initial guess. In all solutions, the value of the last term in c is much smaller than the value of the first two terms.

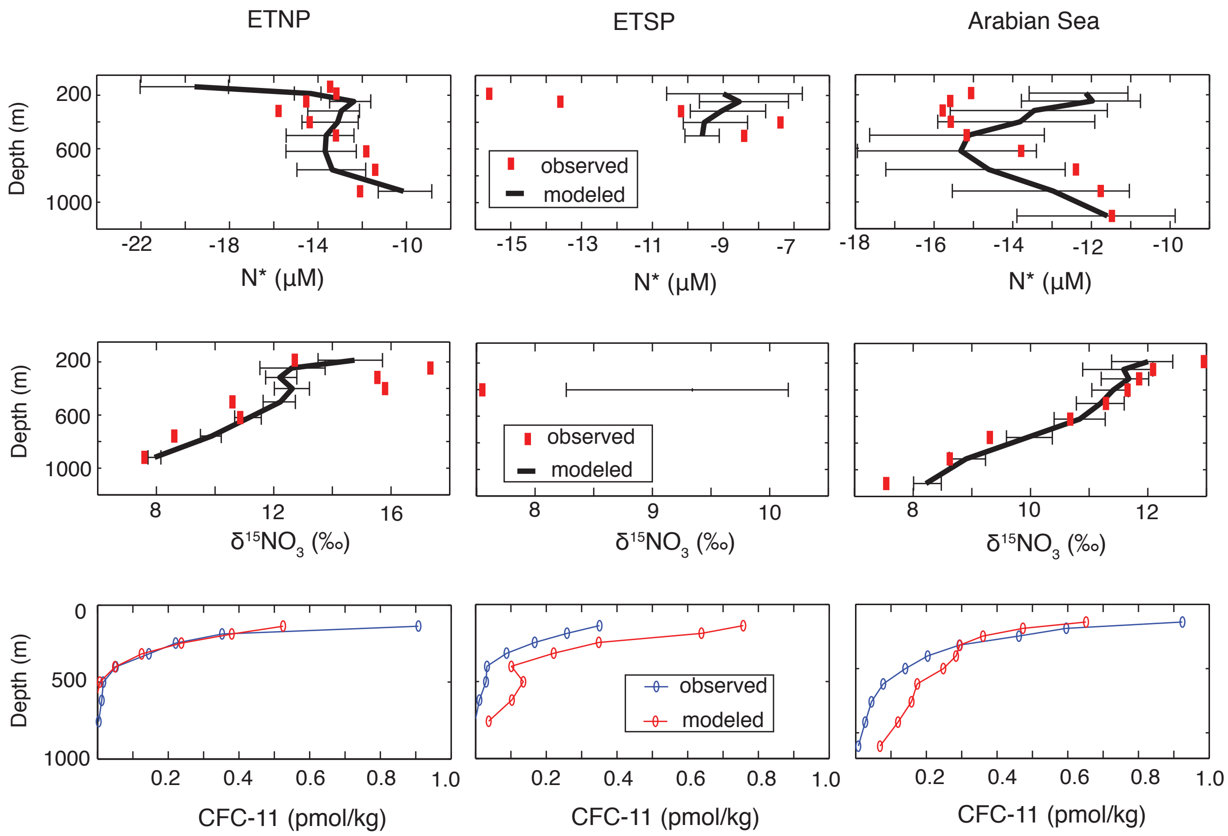

Figure A1: Observed and modeled tracer distributions in the three main suboxic zones: (top) N*, (middle) $\delta^{15}NO_3$, (bottom) CFC-11.