Peer Review File

Global upper ocean dissolved oxygen budget for constraining the biological carbon pump

Corresponding Author: Dr Ryohei Yamaguchi

This file contains all editorial decision letters in order by version, followed by all author rebuttals in order by version.

Attachments originally included by the reviewers as part of their assessment can be found at the end of this file.

Version 0:

Decision Letter:

** Please ensure you delete the link to your author home page in this e-mail if you wish to forward it to your coauthors **

Dear Dr Yamaguchi,

Your manuscript titled "Global upper ocean oxygen budget and an observational constraint on the biological pump" has now been seen by 3 reviewers, whose comments are appended below. You will see that they find your work of some potential interest. However, they have raised quite substantial concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication, but would be interested in considering a revised version that fully addresses these serious concerns.

We hope you will find the reviewers' comments useful as you decide how to proceed. Should additional work allow you to address these criticisms and meet the editorial thresholds outlined below, we would be happy to look at a substantially revised manuscript.

In the following, we list our editorial thresholds based on reviewer comments:

1) Provide a compelling and robust estimate of global annual net community production in the upper ocean.
2) Provide robust uncertainty estimates.
3) Apply a variable depth range that more accurately represents the euphotic zone across the global ocean.
4) Provide an in-depth discussion of the related literature and clearly contextualise and communicate the advance your findings offer to our understanding.

If you choose to take up this option, please either highlight all changes in the manuscript text file, or provide a list of the changes to the manuscript with your responses to the reviewers.

Please bear in mind that we will be reluctant to approach the reviewers again in the absence of substantial revisions.

If the revision process takes significantly longer than three months, we will be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the reviewers’ comments with a list of your changes to the manuscript text (which should be in a separate document to any cover letter), a tracked-changes version of the manuscript (as a PDF file) and any completed checklist:

Link Redacted

** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

Please do not hesitate to contact us if you have any questions or would like to discuss the required revisions further. Thank you for the opportunity to review your work.

Best regards,

Clare Davis, PhD
Senior Editor
Communications Earth & Environment

Michael Stukel, PhD
Editorial Board Member
Communications Earth & Environment
orcid.org/0000-0002-7696-6739

EDITORIAL POLICIES AND FORMAT

If you decide to resubmit your paper, please ensure that your manuscript complies with our editorial policies and complete and upload the checklist below as a Related Manuscript file type with the revised article:

Editorial Policy Policy requirements (Download the link to your computer as a PDF.)

For your information, you can find some guidance regarding format requirements summarized on the following checklist: (https://www.nature.com/documents/commsj-phys-style-formatting-checklist-article.pdf) and formatting guide (https://www.nature.com/documents/commsj-phys-style-formatting-guide-accept.pdf).

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

The authors have taken on an important project with the work in this manuscript. They use the complete BGC-Argo dissolved oxygen data set to compute global ocean annual net community production (ANCP). Understanding the effects of a changing climate on the global carbon cycle requires robust estimates of ANCP at the global scale. The present analysis is impressive in that regard. However, there is one significant drawback that leads me to suggest that the results presented for ANCP have a systematic error. The work needs to be revised to correct what I believe is a misconception.

The authors have computed ANCP using a depth range defined as (lines 108-110), “We defined the depth H of the upper ocean used in the calculation of the upper ocean oxygen budget (equation [1]) as the deeper of the annual maximum mixed layer depth (MLD) or the annual maximum euphotic layer depth (Zeu)”. In regions such as the Southern Ocean and North Atlantic, the annual maximum MLD is more than 400 m (Extended Data Fig. 2b). The depth over which ANCP is computed is 5 times, or more, deeper than the maximum euphotic zone depth (Extended Data Fig. 2d).

The ANCP computed in this manner will include not just production in the euphotic zone, but much of the respiration that occurs below the euphotic zone. ANCP is intended to represent the annual production in the upper ocean, primarily the euphotic zone. The results in this manuscript are essentially an ANCP that includes much of the carbon that is respired below the euphotic zone in the Southern Ocean and North Atlantic. That isn’t a very useful value, in my opinion. It would not, for example, approximate the amount of carbon exported from the surface. In the Southern Ocean and North Atlantic, it would be closer to the amount of carbon sinking at a depth of 400 (or more) m. Since much of the carbon is respired between the euphotic zone and 400 m, it would not be the conventional ANCP.

As a result of this unique ANCP definition, the global ANCP reported in the manuscript of 6 Pg C/yr would be biased low. A very comparable analysis has been done by Wang et al. (2022) using all of the ship-based oxygen data in World Ocean Atlas. The Wang study differs, however, as the depth for ANCP is variable (“using a variable mixed layer depth as in this study”, Wang pg. 10). During summer, they use a much shallower limit for their assessment of the effect of plankton productivity on the oxygen distribution. They find an ANCP of 10 Pg C/y. I find their approach and results more consistent with our oceanographic understanding. The paper by Wang et al. is not cited in the references, but ought to be.

The ANCP values for the Southern Ocean, derived with the constant and very deep depth limit for integration, yield negative results south of about 45 S (Figure 2f). That is inconsistent with nearly all of the prior studies listed in the comprehensive supplementary table. The authors address this discrepancy on page 11 by suggesting that contributions of horizontal and vertical physical processes have been neglected in prior studies. That is not completely correct. Beginning with the study by MacCready and Quay (2001), a variety of estimates of ANCP with positive values include assessments of transport processes. A more likely explanation for the negative ANCP values in the present study is the very deep integration depth used to generate the ANCP values.

The air-sea exchange of oxygen has a major influence on the inferred ANCP, as noted in the paper. It would be helpful to have seen a bit more assessment of the computed air-sea fluxes than the statement on lines 129/130. This might include a comparison of zonal mean fluxes with other estimates such as Gruber et al. (2002) at the global scale or Bushinsky et al. (2017) in the Southern Ocean. This could be part of the supplemental material.

In summary, this manuscript represents a considerable amount of work and it could be a significant contribution to our understanding of global ANCP. However, the results suffer from an inappropriate definition of ANCP and, as currently reported, the results are not very useful. It could be a very valuable contribution if the authors used a more realistic approach to compute ANCP.

Bushinsky, S.M., Gray, A.R., Johnson, K.S. and Sarmiento, J.L., 2017. Oxygen in the Southern Ocean from Argo floats: Determination of processes driving air-sea fluxes. Journal of Geophysical Research: Oceans, 122(11), pp.8661-8682.

Gruber, N., Gloor, M., Fan, S.M. and Sarmiento, J.L., 2001. Air-sea flux of oxygen estimated from bulk data: Implications for the marine and atmospheric oxygen cycles. Global Biogeochemical Cycles, 15(4), pp.783-803.

MacCready, P., and P. Quay (2001), Biological export flux in the Southern Ocean estimated from a climatological nitrate budget, Deep Sea Res., Part II, 48, 4299–4322, doi:10.1016/S0967-0645(01)00090-X.

Wang Z, Garcia HE, Boyer TP, Reagan J and Cebrian J (2022) Controlling factors of the
climatological annual cycle of the surface mixed layer oxygen content: A global view.
Front. Mar. Sci. 9:1001095. doi: 10.3389/fmars.2022

Reviewer #2 (Remarks to the Author):

The author’s present a global synthesis of dissolved surface ocean oxygen data to estimate the rate at which organic matter is transferred (exported) from the surface ocean to deeper depths. This process, called the ocean’s biological pump, is a key factor in the ocean’s ability to sequester CO2 from the atmosphere and distribution of nutrients and O2 in the deep ocean. It is important to improve our current state of knowledge of the biological pump in order to detect and predict future changes in response to ocean warming. From this perspective, the paper significantly contributes towards improving our understanding of regional variations in the strength of the ocean’s biological pump.

There are some significant issues that need to be resolved, however, before I can recommend publication of the paper, as described below.

Overall Issues
1. Error Analysis (Table 1):
The estimated errors in the air-sea and vertical diffusive O2 fluxes seem unreasonably low.

The uncertainty in the diffusive component of air-sea O2 flux has been estimated at ±25% and the bubble component at ±10% which yields errors of ±0.2 to 0.5 mol C/m2/yr for annual NCP (e.g., Yang et al, 2017) which is equivalent to ±104 - 260 Tmol O2/yr. This error range greatly exceeds the errors of ±33 and ±7 Tmol O2/yr reported in Table 1.

The reported uncertainty in the vertical diffusive O2 flux is ~±25% of the flux (Table 1). Yet when float or mooring based T and S budgets are used to estimate Kz values at the base of the mixed layer the typical errors range from order ±50% during the summer and up to 10-fold in winter (see Cronin et al., JGR, 2015). Thus, the authors choices of three Kz values doesn’t seem to adequately represent uncertainty error in Kz and thus vertical O2 flux. Nor is it clear how the authors chose between the three Kz values at specific grid points.

Because the same air-sea O2 gas transfer equation is used at each grid point the errors in the air-sea O2 flux at each grid point are not independent of each other. Thus, the error in the mean air-sea O2 flux determined for a region is not reduced by the (square root) number of estimates. A Monte Carlo approach has often been used to account for uncertainties in O2 budget based estimates of export (e.g., Yang et al, 2017) but this approach may be computationally too time consuming to utilize with a global grid.

2. Comparison to Quay’s paper ‘Organic Matter Export Rates and the Pathways of Nutrient Supply in the Ocean’ which appeared in Global Biogeochemical Cycles in 2023.

The authors should compare their results to those presented in Quay 2023 which estimated export (ANCP) based on surface O2 saturation on a global scale. Using World Ocean Atlas surface O2 saturation data to estimate air-sea O2 flux (but ignoring physical transport terms) Quay estimated a global ANCP of 2.2±0.8 mol C/m2/yr between 50°S and 50°N. This estimate is equivalent to ~800 Tmol O2/yr which is more than double the air-sea O2 flux of 321 Tmol O2/yr (north of 45 °S) reported in Table 1. Since both studies used global O2 databases and the same air-sea bubble and diffusive O2 gas transfer calculation (Emerson et al., 2019), the author’s should provide some possible reasons why these two estimates of air-sea O2 flux are so different. Furthermore, the difference between these two estimates underscores the point about error analysis discussed above, i.e., is the reported uncertainty in the global air-sea O2 flux of ±7 Tmol O2/yr (Table 1) realistic?

On the positive side, the results of the present study share some similar conclusions to Quay (2023), i.e., highest ANCP in N. Atlantic, higher ANCP in northern versus southern basins of Pacific and Atlantic and lowest ANCP in Indian Ocean. Theses similarities are worth pointing out.

3. O2 budgets in regions of intermediate and deep water formation.

The 7-fold difference between the estimated air-sea O2 flux of -46 TmolO2/yr for global ocean versus 321 T mol O2/yr for the ocean north of 45°S underscores the sensitivity of the author's global O2 budget to the processes occurring in the southern ocean. The undersaturated surface O2 levels south of ~45°S (Fig 1a) are a result of upwelling and deep mixing bringing up O2 deficient water. Since almost all of the surface waters in this region subduct below 300m or so via Subtropical and SubAntarctic Mode Water and Antarctic Intermediate Water (south of 35°S) they participate in the O2 budget of the subsurface ocean rather than the surface ocean. A similar situation occurs in the polar N. Atlantic (>60°N) during deep convection related to North Atlantic Deep Water formation. Thus to include these intermediate and deep water formation regions when calculating the O2 budget of the "surface" ocean seems to substantially deepen the effective depth horizon of the O2 budget and, as a result, likely yields an underestimated of the global ANCP for the 'surface' ocean. The author's should explain how this situation affects their results.

4. Comparison to other model-based estimates of export

Previous studies have taken a similar approach as the author's by using a model combined with observations to estimate global export. Devries et al (Glob. Biogeo. Chem., 2017) used a model that fit satellite productivity and observed particle sinking rate, dissolved organic carbon and subsurface O2 and estimated a global export rate of 9 Gt C/yr at the base of euphotic layer. Schlitzer (J. Oceanogr., 2004) used a global circulation model and observed nutrient and O2 fields to estimate a global export rate of 9.6 Gt C/yr at the base of the euphotic layer. The author's need to compare their results to these previous results and offer possible explanations of why their surface O2 based export estimate of 6 GtC/yr is significantly lower than these previous estimates.

5. Disagreement with observations south of 50° S

Although there is good agreement between the author's results with previous estimates north of 40°S their estimated ANCP is consistently lower than previous estimates south of 40°S. The author's should state possible explanations for this discrepancy.

Specific issues

Line 425: Did they include spatial and temporal variations in atmospheric pressure when calculating O2 saturation state which is particularly important south of 50°S. The atmospheric pressure term should be included in equation 5.

Line 453: Although they state the use of three different Kz values (4,6 and 8 e-4 m2/s) at the base of the mixed layer to estimate the vertical flux of O2, it wasn't clear how they choose which Kz value to use at a specific location and time.

Line 510: They should state the fraction of the grid that was missing due to lack of data.

Reviewed by Paul Quay

Reviewer #3 (Remarks to the Author):

Please see attached pdf.

For more information please visit http://www.springernature.com/orcid

If you experience problems in linking your ORCID, please contact the Platform Support Helpdesk.

Version 1:

Decision Letter:

** Please ensure you delete the link to your author home page in this e-mail if you wish to forward it to your coauthors **

Dear Dr Yamaguchi,

Please allow us to apologise for the delay in sending a decision on your manuscript titled "Global upper ocean dissolved oxygen budget for constraining the biological carbon pump". It has now been seen by our reviewers, whose comments appear below. In light of their advice we are delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment, provided you discuss in the manuscript and account for the issue regarding the definition of ANCP in high latitude regions raised by Reviewer #1.

We therefore invite you to revise your paper one last time to address the remaining concerns of our reviewers. At the same time we ask that you edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table".

*****Please take care to match our formatting and policy requirements. We will check revised manuscript and return manuscripts that do not comply. Such requests will lead to delays. *****

Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files as a Related Manuscript file.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at https://www.nature.com/documents/commsj-file-checklist.pdf .

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication. For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit https://www.nature.com/commsenv/open-access

At acceptance, you will be provided with instructions for completing the open access licence agreement on behalf of all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:
Link Redacted
** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Michael Stukel, PhD
Editorial Board Member
Communications Earth & Environment
orcid.org/0000-0002-7696-6739

Joe Aslin
Deputy Editor,
Communications Earth & Environment
https://www.nature.com/commsenv/
Twitter: @CommsEarth

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

The authors have generally done a nice job of responding to the reviews, particularly to comments about uncertainty and statistics. It is a very nice piece of work and has global significance. However, I will make one very critical comment that seem easily fixable. That regards my comment with respect to their decision to calculate Annual Net Community Production by integrating to the maximum winter mixed layer depth. That decision leads to a systematic bias low in their results because the mass balance they use includes significant amounts of respiration below the euphotic zone. They acknowledge that “our ANCPs are on the smaller side” on lines 270/273 due to the “greater influence of subsurface respiration” and this is clearly shown in Figure S7a where they compare their results to other studies. It shows a bias low of 1/3, which I believe is a new addition to the supplementary material.

Annual net community production (ANCP) is defined as the difference between net primary production (NPP) and respiration by heterotrophs (animals and bacteria) in the upper ocean. ANCP is also equated to the amount of carbon that is exported from the upper ocean over an annual cycle. For example, as stated in Emerson and Hedges (Chemical Oceanography and the Marine Carbon Cycle, 2008, pg. 30), “the mean organic carbon flux out of the euphotic zone should equal the net community production of organic carbon”. Note that this clearly defines the upper ocean in the determination of NCP as the euphotic zone. Unfortunately there are a number of studies of ANCP, mostly based on tracer budgets, that have used the maximum depth of the mixed layer to define ANCP. That happens for reasons explained in this manuscript (one doesn’t have to calculate entrainment and detrainment). These studies primarily occur at places such as Ocean Station Papa or the Hawaii Ocean Time-series site, where the maximum mixed layer depth is not particularly deep and not strikingly different than the euphotic zone. So the biases aren’t large in these other studies and have largely eluded comment. However, in the case of the global analysis in this manuscript, much of the ocean has quite deep mixed layers. Figure 3 shows everything poleward of 40 degrees as deeper than 200 m and well below the euphotic zone. No wonder their global ANCP seems low.

Does this mean that the results of the paper are invalid? No, it’s a really nice piece of work. The authors could fix it simply by focusing on their statement that their results are “the equivalent amount of carbon, in either particulate or dissolved form, escapes from the predominant upper-ocean seasonal cycle of physical environmental conditions (e.g., mixing and restratification) and becomes sequestered into the ocean interior on longer timescales than, at shortest, one year.” (lines 202/204). That is a critical value for our understanding of the global carbon cycle. Rather than calling the results ANCP, maybe use something like Annual Carbon Sequestration? They could get credit for coining the term. They could clearly note that their method is highly related to ANCP and yields a lower limit. But they should avoid trying to equate it to ANCP.

Reviewer #3 (Remarks to the Author):

Thank you to the authors for their comprehensive responses to the reviewer suggestions. I think the current version of the manuscript has sufficiently responded to all comments and is ready for publication. I would also like to add that I think their use of the seasonal maximum mixed layer depth for their ANCP calculation makes sense and is the correct approach for this type of work, given their research questions.

Seth Bushinsky

Reviewer #1

The authors have taken on an important project with the work in this manuscript. They use the complete BGC-Argo dissolved oxygen data set to compute global ocean annual net community production (ANCP). Understanding the effects of a changing climate on the global carbon cycle requires robust estimates of ANCP at the global scale. The present analysis is impressive in that regard. However, there is one significant drawback that leads me to suggest that the results presented for ANCP have a systematic error. The work needs to be revised to correct what I believe is a misconception.

Thank you for dedicating time to carefully review our manuscript. We have incorporated the following revisions and additions in response to your comments:

1) We have provided an explanation (L106–121, and L473–485) and added relevant discussions (L261–271, L322–335, Note S3, and Fig. S7) regarding the definition of the upper ocean box in the oxygen budget calculation.
2) We have revised the method for calculating the air-sea oxygen flux incorporating a comment from Reviewer #2 (inclusion of sea surface pressure correction). And, we have included comparisons of the new result with previous study you pointed and associated discussions on air-sea oxygen fluxes (Note S2 and Fig. S6).

Please refer to our responses to each paragraph for details.

The authors have computed ANCP using a depth range defined as (lines 108-110), “We defined the depth H of the upper ocean used in the calculation of the upper ocean oxygen budget (equation [1]) as the deeper of the annual maximum mixed layer depth (MLD) or the annual maximum euphotic layer depth (Zeu)”. In regions such as the Southern Ocean and North Atlantic, the annual maximum MLD is more than 400 m (Extended Data Fig. 2b). The depth over which ANCP is computed is 5 times, or more, deeper than the maximum euphotic zone depth (Extended Data Fig. 2d).
The ANCP computed in this manner will include not just production in the euphotic zone, but much of the respiration that occurs below the euphotic zone. ANCP is intended to represent the annual production in the upper ocean, primarily the euphotic zone. The results in this manuscript are essentially an ANCP that includes much of the carbon that is respired below the euphotic zone in the Southern Ocean and North Atlantic. That isn’t a very useful value, in my opinion. It would not, for example, approximate the amount of

carbon exported from the surface. In the Southern Ocean and North Atlantic, it would be closer to the amount of carbon sinking at a depth of 400 (or more) m. Since much of the carbon is respired between the euphotic zone and 400 m, it would not be the conventional ANCP.

Thank you for your comments. We think that your interpretation of our results is correct, but we also believe that it is not necessarily required to define the lower boundary of the upper ocean in the dissolved oxygen budget calculation (i.e., the depth at which an ANCP is estimated, referred to as depth *H* in this study) strictly as the base of the euphotic zone. Since an NCP is simply determined as a net primary production (NPP) minus heterotrophic respiration (HR), in principle, it can be defined at any depth. As you pointed out, however, care must be taken in defining the depth plane to use obtained ANCP value as an approximation of "carbon exported from the layer above" under an assumption of steady state. It is necessary to properly define the *H*-plane along which carbon export is to be estimated.

In regions where the annual maximum mixed layer depth (MLD) exceeds the annual maximum euphotic layer depth (Z_{eu}), the majority of the organic carbon exported from the euphotic layer once during warm productive season (when the mixed layer is shallow) is re-entrained into the euphotic layer during winter mixing and then respired and reused for production (Palevsky and Quay 2017; Wang et al. 2023). The objective of this study is to examine the carbon export at the depth of the winter mixed layer, where the sequestration timescales are expected to be at least one year. And, it contrasts with the carbon export that is expected to be involved again in the mixed and/or euphotic layer within a year. We have clarified this background of our study (L106–121), also have discussed the effect of the respiration below the euphotic zone in the result section (L261–271, and L322–335).

In addition to the scientific rationale based on our scientific motivation outlined above, there are also practical considerations that justify the utilization of the definition of depth *H* used in our study. By employing a time-invariant depth as the lower boundary of the upper ocean box for budget calculation, it is possible to exclude the entrainment and detrainment terms associated with the change in the thickness of the box over time. It is generally difficult to accurately estimate the entrainment/detrainment terms in any tracer budget calculation using low temporal resolution data (cf. Kim et al. 2006). At the time of this analysis and still today, monthly was the highest temporal resolution available globally for dissolved oxygen budget calculation. Moreover, our definition of H indicates that the surface mixed layer is always within the box throughout the year. This is not achieved using the Z_{eu} as the depth H. Consequently, the inclusion of several orders of magnitude larger diffusion within the mixed layer, which would introduce considerable errors, is not necessary in the budget calculation. Thus, while imperfect, we believe that this definition provides the most robust method to date for obtaining an ANCP estimate from data with current limited spatiotemporal coverage. This practical consideration has clarified in the method section (L473–485).

As a result of this unique ANCP definition, the global ANCP reported in the manuscript of 6 Pg C/yr would be biased low. A very comparable analysis has been done by Wang et al. (2022) using all of the ship-based oxygen data in World Ocean Atlas. The Wang study differs, however, as the depth for ANCP is variable ("using a variable mixed layer depth as in this study", Wang pg. 10). During summer, they use a much shallower limit for their assessment of the effect of plankton productivity on the oxygen distribution. They find an ANCP of 10 Pg C/y. I find their approach and results more consistent with our oceanographic understanding. The paper by Wang et al. is not cited in the references, but ought to be.

A number of previous studies share similar scientific motivation mentioned above and use the same definition of H (annual maximum MLD). Therefore, we believe that our definition of H is certainly not "unique". This definition is used not only in studies estimating ANCP from the tracer budgets (studies labeled as WMLD in Table S1) but also in studies directly quantifying ocean carbon cycling (e.g., Levy et al. 2013). This is likely because their results can be interpreted easily and give implications on the upper ocean carbon cycle. It has become crucial, for further understanding the role of the biological pump on the climate system, to assess not just the quantification of the export from the upper ocean but also to assess the duration for which the exported carbon remains isolated from the air-sea interface. In that context, our estimate can be interpreted as global export that has its sequestration timescale of, at shortest, one-year. Our revised estimate of global carbon export, with the expected sequestration timescale of at least one year, is consistent with the latest independent result obtained by an inverse model (Wang et al. 2023), indicating that our estimates are not necessarily "biased low". We have added the comparisons with previous estimates (Note S3, and Fig. S7) and associated discussion in the main text (L261–271).

When using a time-varying seasonal mixed layer depth (MLD) for the dissolved oxygen budget (ANCP calculation), as Wang et al. (2022) have done, the resulting estimate of ANCP does not encompass the production that occurs below the summertime thin mixed layer but within the euphotic zone (e.g., the subsurface oxygen maximum in subtropics). Moreover, as previously mentioned, calculating the entrainment/detrainment term from monthly data can introduce nonnegligible errors. Beyond the definition of the upper ocean box, there are several other differences between our study and Wang et al. (2022): 1) All physical oxygen fluxes except for air-sea exchange and entrainment are ignored, which are also important for closing the global upper ocean dissolved oxygen balance (Fig. 3 and Fig.4b). 2) Using monthly mean gridded dissolved oxygen data for the nonlinear air-sea flux calculations does not capture high-frequency components of sea surface dissolved oxygen variability. 3) The flux calculation does not include a sea level pressure correction and a skin layer temperature correction, which significantly impacts both the global integration and its spatial pattern (Note S2 and Fig. S6). Therefore, it is difficult to compare these results directly. But, on the positive side, the definition of *H* in Wang et al. (2022), which should be shallower than ours on a global mean basis and thus is expected to reflect more upper layer production, naturally results in a larger estimate of global ANCP than our estimates. We have added the paper in the References (L269).

The ANCP values for the Southern Ocean, derived with the constant and very deep depth limit for integration, yield negative results south of about 45 S (Figure 2f). That is inconsistent with nearly all of the prior studies listed in the comprehensive supplementary table. The authors address this discrepancy on page 11 by suggesting that contributions of horizontal and vertical physical processes have been neglected in prior studies. That is not completely correct. Beginning with the study by MacCready and Quay (2001), a variety of estimates of ANCP with positive values include assessments of transport processes. A more likely explanation for the negative ANCP values in the present study is the very deep integration depth used to generate the ANCP values.

Regarding the remarkable discrepancies in the air–sea flux estimates with previous studies over the Southern Ocean in the initial manuscript, the results have been improved by incorporating the comments from Reviewer #2 (implementation of a sea level pressure correction in the air-sea oxygen calculation). We have deleted the discussion of the discrepancy in the Southern Ocean. And, comparisons with previous studies, including

MacCready and Quay (2001), are shown in Figure 4, Note S3, and Figure S7.

The air-sea exchange of oxygen has a major influence on the inferred ANCP, as noted in the paper. It would be helpful to have seen a bit more assessment of the computed air-sea fluxes than the statement on lines 129/130. This might include a comparison of zonal mean fluxes with other estimates such as Gruber et al. (2002) at the global scale or Bushinsky et al. (2017) in the Southern Ocean. This could be part of the supplemental material.

The supplementary information (Fig. S6 and Note S2) includes a comparison of air-sea flux with Gruber et al. (2001) and Bushinsky et al. (2017). The data source (BGC Argo data) and air-sea flux parameterizations were shared with Bushinsky et al. (2017), and it was confirmed that the result in the Southern Ocean was consistent (gray and blue bars in Fig. S4). A comparison with Gruber et al. (2001) revealed that while there are some consistent regions, there are still considerable discrepancies in the estimated air-sea oxygen flux from different methods, particularly in the mid-to-high latitudes of the North Hemisphere. The air-sea fluxes exhibit the highest uncertainty in the dissolved ocean budget framework (Fig. 3), indicating the necessity for further studies to address these issues in the future.

In summary, this manuscript represents a considerable amount of work and it could be a significant contribution to our understanding of global ANCP. However, the results suffer from an inappropriate definition of ANCP and, as currently reported, the results are not very useful. It could be a very valuable contribution if the authors used a more realistic approach to compute ANCP.

Thank you for carefully reading our initial manuscript. We think that your concern is valid and we totally agree your concern. While we believe that the ability to estimate NCPs at any depth, including the euphotic zone, mixed layer, and seasonal thermocline throughout the year, would contribute to a more detailed understanding of the mechanisms of biological processes occurring in the upper ocean, we think that a more spatiotemporally dense observational data or integrated method with model is needed to do so globally. Therefore, we will leave that challenge as a task for future work and it is expected to be able to address it as local campaign.

In this study, because of our scientific motivation (estimation of carbon export which has the sequestration timescale of at least one year), we have decided not to change the definition of H in the revised manuscript. Instead, we have added the explanation in the introduction and included a discussion about the definition of H as listed above.

We would be grateful if you would kindly undertake a further review of our work.

References

Bushinsky, S. M., A. R. Gray, K. S. Johnson, and J. L. Sarmiento, 2017: Oxygen in the Southern Ocean From Argo Floats: Determination of Processes Driving Air-Sea Fluxes. *J. Geophys. Res. Oceans*, **122**, 8661–8682, https://doi.org/10.1002/2017JC012923.

Gruber, N., M. Gloor, S.-M. Fan, and J. L. Sarmiento, 2001: Air-sea flux of oxygen estimated from bulk data: Implications For the marine and atmospheric oxygen cycles. *Global Biogeochemical Cycles*, **15**, 783–803, https://doi.org/10.1029/2000GB001302.

Kim, S.-B., I. Fukumori, and T. Lee, 2006: The Closure of the Ocean Mixed Layer Temperature Budget Using Level-Coordinate Model Fields. *Journal of Atmospheric and Oceanic Technology*, **23**, 840–853, https://doi.org/10.1175/JTECH1883.1.

Levy, M., L. Bopp, P. Karleskind, L. Resplandy, C. Ethe, and F. Pinsard, 2013: Physical pathways for carbon transfers between the surface mixed layer and the ocean interior: PHYSICAL CARBON FLUXES. *Global Biogeochem. Cycles*, **27**, 1001–1012, https://doi.org/10.1002/gbc.20092.

MacCready, P., and P. Quay, 2001: Biological export flux in the Southern Ocean estimated from a climatological nitrate budget. *Deep Sea Research Part II: Topical Studies in Oceanography*, **48**, 4299–4322, https://doi.org/10.1016/S0967-0645(01)00090-X.

Palevsky, H. I., and P. D. Quay, 2017: Influence of biological carbon export on ocean carbon uptake over the annual cycle across the North Pacific Ocean: Influences on North Pacific Ocean CO $_2$ Uptake. *Global Biogeochem. Cycles*, **31**, 81–95, https://doi.org/10.1002/2016GB005527.

Wang, W.-L., W. Fu, F. A. C. Le Moigne, R. T. Letscher, Y. Liu, J.-M. Tang, and F. W. Primeau, 2023: Biological carbon pump estimate based on multidecadal hydrographic data. *Nature*, **624**, 579–585, https://doi.org/10.1038/s41586-023-06772-4.

Wang, Z., H. E. Garcia, T. P. Boyer, J. Reagan, and J. Cebrian, 2022: Controlling factors of the climatological annual cycle of the surface mixed layer oxygen content: A global view. *Frontiers in Marine Science*, **9**.

Reviewer #2

The author's present a global synthesis of dissolved surface ocean oxygen data to estimate the rate at which organic matter is transferred (exported) from the surface ocean to deeper depths. This process, called the ocean's biological pump, is a key factor in the ocean's ability to sequester CO2 from the atmosphere and distribution of nutrients and O2 in the deep ocean. It is important to improve our current state of knowledge of the biological pump in order to detect and predict future changes in response to ocean warming. From this perspective, the paper significantly contributes towards improving our understanding of regional variations in the strength of the ocean's biological pump.
There are some significant issues that need to be resolved, however, before I can recommend publication of the paper, as described below.

Thank you for taking the time to carefully read our manuscript. In response to your comments, we have made the following revisions:

(1) We recalculated the errors for the terms derived from individual profiles in the dissolved oxygen budget calculation (air-sea flux and vertical diffusion) using a Monte Carlo approach. As the result, the error for the air-sea flux term estimated appropriately larger, while the error for the diffusion term slightly decreased (Fig. 3, and Method).
(2) We introduced a correction for changes in saturated oxygen concentration due to atmospheric pressure in the calculation of the air-sea flux term. This adjustment resulted in reduced oxygen uptake in the Southern Ocean, where sea level pressure is generally low throughout the year, making the latitudinal distribution of ANCP more consistent with previous studies (Fig. S6, and Fig. 4).
(3) According to the changes in the results, we revised related discussions (Section "Global ocean ANCP based on dissolved oxygen budget (L217–)" and "Implications for the global ocean carbon cycle (L354–)").

For more detailed points and other minor changes, please refer to our responses to individual comments.

Overall Issues
1. Error Analysis (Table 1):
The estimated errors in the air-sea and vertical diffusive O2 fluxes seem unreasonably

low.

The uncertainty in the diffusive component of air-sea O2 flux has been estimated at ±25% and the bubble component at ±10% which yields errors of ±0.2 to 0.5 mol C/m2/yr for annual NCP (e.g., Yang et al, 2017) which is equivalent to ±104 - 260 Tmol O2/yr. This error range greatly exceeds the errors of ±33 and ±7 Tmol O2/yr reported in Table 1.

Thank you for your comments. As you and Reviewer #3 pointed out, we recognize that the error estimation in the first manuscript was inadequate. In the revised manuscript, we have changed the error estimation method, primarily following the approach of Bushinsky et al. (2017).

In the calculation of physical oxygen fluxes derived from individual profiles (air–sea exchange and vertical diffusion), we used a Monte Carlo approach to quantify associated uncertainties. In the air–sea flux calculation, for each profile, the flux calculations were repeated 1,500 times using randomly selected values. A randomly selected wind value was taken from a normal distribution with a mean of an average value of five spatially interpolated daily wind products and a standard deviation of five values. The surface oxygen concentration was also randomly taken from a normal distribution with the observed surface oxygen as the mean and the 1% error of the mean as the standard deviation. The mean of the iterations was taken as the flux value, and its standard deviation was employed as a measure of uncertainty. Flux values and their errors, obtained from 1,500 repeated calculations across ~ 600,000 profiles, were spatially mapped to a 1° (longitude) x 1° (latitude) grid. The errors associated with the globally integrated fluxes are reported as the simple summation of errors mapped.

As the result, The error for the global integrated air-sea oxygen exchange (206 Tmol O_2 year^{-1}) was estimated to be 246 Tmol O_2 year^{-1}, yielding resultant global mean annual net community production (ANCP) of 1.7 ± 0.5 mol C m^{-2} year^{-1}. According to these changes, we have revised the result section (Fig. 3 and 4, L196–200, L218–231) and "Uncertainty estimation" in Methods (L588–617).

The reported uncertainty in the vertical diffusive O2 flux is ~±25% of the flux (Table 1). Yet when float or mooring based T and S budgets are used to estimate Kz values at the base of the mixed layer the typical errors range from order ±50% during the summer and up to 10-fold in winter (see Cronin et al., JGR, 2015). Thus, the authors choices of three Kz values doesn't seem to adequately represent uncertainty error in Kz and thus vertical O2 flux. Nor is it clear how the authors chose between the three Kz values at specific grid points.

We apologize for any confusion regarding the vertical diffusion coefficients used. In the first manuscript, we assumed three values for the vertical diffusion coefficient at the mixed layer base (common globally and time-invariant) and calculated the vertical diffusion flux globally for each case (three cases) by combining them with two gridded products of the background vertical diffusion coefficient (resulting in six cases). We then calculated the mean of the six results and estimated the uncertainty from their standard deviation. The three values were annual means of results from Cronin et al. (2015).

In the revised manuscript, we used the monthly vertical diffusion coefficients at the mixed layer base and their reported uncertainties (~100 % error) from Cronin et al. (2015). As the monthly vertical diffusion coefficients were derived from observations in the Northern Hemisphere, they were applied directly to the grid points of the Northern Hemisphere (north of 20ºN). In the tropical grid points (20°S–20°N), the annual mean value was used throughout the year. In the Southern Hemisphere (south of 20°S), a six-month shifted monthly coefficient was used. We have clarified the explanation for calculation of the vertical diffusivity (L518–533).

With regard to the uncertainty estimate of the vertical diffusion flux, a Monte Carlo approach was employed in the revised manuscript in a manner analogous to the air-sea flux calculation. For each profile, the vertical diffusion flux calculations were repeated 1,500 times using randomly selected values for the background and mixed layer base diffusion coefficients. The randomly selected background diffusivity was derived from two independent sources, while the diffusivity at the base of the mixed layer was assumed to have a 100% error relative to the monthly diffusion coefficients from Cronin et al. (2015). The mean of these iterations was taken as the flux value, and its standard deviation was employed as a measure of uncertainty. Consequently, the estimated error for the global integrated vertical diffusion flux (248 Tmol O_2 year^{-1}) was determined to be 33 Tmol O_2 year^{-1}. In comparison to the air-sea flux, the smaller uncertainty in the global integrated value can be attributed to the fact that the air-sea flux is relatively large in most regions and months, while vertical diffusion is effective only in specific regions and only during winter. This tendency, where the uncertainty in the vertical diffusion term is less pronounced compared to the air-sea flux uncertainty, is consistent with previous studies (e.g., Yang et al. 2017), which were conducted in regions where the vertical diffusion is comparatively important in the budget. According to these changes, we have revised the Methods “Uncertainty estimation” (L588–617)

Because the same air-sea O2 gas transfer equation is used at each grid point the errors in the air-sea O2 flux at each grid point are not independent of each other. Thus, the error in the mean air-sea O2 flux determined for a region is not reduced by the (square root) number of estimates. A Monte Carlo approach has often been used to account for uncertainties in O2 budget based estimates of export (e.g., Yang et al, 2017) but this approach may be computationally too time consuming to utilize with a global grid.

Thank you again for your valuable comments. We believe that the recalculations using the Monte Carlo approach in the revised manuscript provide more reasonable error estimates.

2. Comparison to Quay’s paper ‘Organic Matter Export Rates and the Pathways of Nutrient Supply in the Ocean’ which appeared in Global Biogeochemical Cycles in 2023. The authors should compare their results to those presented in Quay 2023 which estimated export (ANCP) based on surface O2 saturation on a global scale. Using World Ocean Atlas surface O2 saturation data to estimate air-sea O2 flux (but ignoring physical transport terms) Quay estimated a global ANCP of 2.2±0.8 mol C/m2/yr between 50°S and 50°N. This estimate is equivalent to ~800 Tmol O2/yr which is more than double the air-sea O2 flux of 321 Tmol O2/yr (north of 45 °S) reported in Table 1. Since both studies used global O2 databases and the same air-sea bubble and diffusive O2 gas transfer calculation (Emerson et al., 2019), the author’s should provide some possible reasons why these two estimates of air-sea O2 flux are so different. Furthermore, the difference between these two estimates underscores the point about error analysis discussed above, i.e., is the reported uncertainty in the global air-sea O2 flux of ±7 Tmol O2/yr (Table 1) realistic? On the positive side, the results of the present study share some similar conclusions to Quay (2023), i.e., highest ANCP in N. Atlantic, higher ANCP in northern

versus southern basins of Pacific and Atlantic and lowest ANCP in Indian Ocean. Theses similarities are worth pointing out.

First of all, because Quay (2023) does not account for advection and horizontal diffusion fluxes in their oxygen budget calculation, described upper ocean dissolved oxygen mass balance is different from that of this study. In Quay (2023), an oxygen loss from the upper ocean due to the air-sea exchange was estimated ~800 Tmol O_2 year^{-1} and an oxygen loss by vertical diffusion was reported as ~10% of the air-sea flux. These losses was balanced by the net biological oxygen production corresponding to an ANCP of 2.2 ± 0.8 mol C m^{-2} year^{-1}. On the other hand, our results (but recalculated in the area aligned with Quay (2023)) demonstrated that the total ~590 Tmol O_2 year^{-1} of oxygen loss, which is the sum of contributions from air-sea exchanges (~270 Tmol O_2 year^{-1}), diffusion (~190 Tmol O_2 year^{-1}), and advection (~130 Tmol O_2 year^{-1}), should be balanced by the net biological production corresponding an ANCP of 1.4 ± 0.5 mol C m^{-2} year^{-1}. Despite the discrepancy between the two studies, it is noteworthy that both ANCP estimates are consistent in terms of the overall characteristic of the spatial pattern, such as hemispheric and basin contrasts, and the two global estimates, including the range of uncertainty, overlap each other. It is likely because that the contributions from processes not considered in Quay (2023), such as advection and horizontal diffusion, affect the global integrated values of ANCP, yet have a limited impact on determining the basin contrast.

A comparison of the globally integrated air-sea oxygen fluxes calculated using similar approaches in both studies reveals that our estimate tends to be less release of oxygen from the ocean compared to the results of Quay (2023). This discrepancy might be attributed to two primary factors: (1) differences in the spatial and temporal resolution of the wind data used, and (2) the inclusion of a skin layer temperature correction. If one assumes that Quay (2023) employed spatiotemporally more smoothed data (e.g., monthly climatological wind), it would likely yield different results from this study. As the parameterization of the air-sea oxygen flux is a nonlinear function of wind speed, the monthly average of fluxes calculated from daily wind speeds is not always identical to the flux calculated from monthly average wind speeds. Indeed, when fluxes were estimated using the same methodology, except for employing climatological mean monthly wind data, it was found that the underestimation of ocean oxygen uptake at high latitudes under strong winds resulted in a globally integrated flux with a 25% increase in ocean oxygen release.

The second reason could be the skin layer temperature correction employed in this study. Recent observational studies have suggested considering the skin effect when calculating the air–sea diffusion fluxes of CO_2 and oxygen (Watson et al. 2020; Yang et al. 2022). Generally, by using a slightly cooler skin layer temperature (temperature at the air–sea interface) instead of the bulk sea surface temperature (temperature at 1¬5-m depth) to estimate the saturation concentration in diffusion flux calculations, the gas solubility increases, leading to a shift in the resultant flux towards more absorption by the ocean. Indeed, the same calculation, excluding the skin layer temperature correction, yielded ~250 Tmol O_2 year^{-1} increase in the ocean oxygen release on a global integrate basis. As previously reported by (Yang et al. 2022), the differences are more pronounced in the subtropical regions (red and purple bars in Fig. S4).

These discussions have been added in Supplementary information (Note S2).

3. O2 budgets in regions of intermediate and deep water formation.
The 7-fold difference between the estimated air-sea O2 flux of -46 TmolO2/yr for global ocean versus 321 T mol O2/yr for the ocean north of 45°S underscores the sensitivity of the author’s global O2 budget to the processes occurring in the southern ocean.

Thanks to your comments (first specific comment regarding se level pressure, discussed later), the revised manuscript significantly improved the estimation of air-sea fluxes, resulting in a somewhat changed overall picture of global ANCP pattern, that is, the relative contribution of the Southern Ocean has decreased. Therefore, the discussion on the differences with previous studies in the Southern Ocean has been removed.

The undersaturated surface O2 levels south of ~45°S (Fig 1a) are a result of upwelling and deep mixing bringing up O2 deficient water. Since almost all of the surface waters in this region subduct below 300m or so via Subtropical and SubAntarctic Mode Water and Antarctic Intermediate Water (south of 35°S) they participate in the O2 budget of the subsurface ocean rather than the surface ocean. A similar situation occurs in the polar N. Atlantic (>60°N) during deep convection related to North Atlantic Deep Water formation. Thus to include these intermediate and deep water formation regions when calculating the O2 budget of the "surface" ocean seems to substantially deepen the effective depth horizon of the O2 budget and, as a result, likely yields an underestimated of the global ANCP for the 'surface' ocean. The author's should explain how this situation affects their results.

We agree with this point. In the deep and intermediate water formation regions of the Southern Ocean and North Atlantic, vertical movements of water masses related to subduction and obduction result in a net loss of oxygen from the upper ocean. In the context of our budget calculations, this effect is accounted for as an advection term, and this effect is indeed observed in these regions (blue areas in Fig. 2c). While we have not overlooked the effect of the advection term itself (i.e., we are not unfairly underestimating the ANCP), as you rightly pointed out, our definition of the lower boundary of the upper ocean may be somewhat too deep to apply the term "surface" to the estimate ANCPs, especially in such deep water formation regions.

On the other hand, we believe there is a certain rationale for selecting the annual maximum mixed layer depth (MLD) as the lower boundary of the upper ocean dissolved oxygen budget (referred to as depth H in the manuscript). As you also mentioned in Quay (2023), the depth horizon of the bottom of the winter mixed layer can be interpreted as the depth limit directly influenced by the pronounced upper ocean seasonal cycle. Therefore, the export flux across this boundary is anticipated to have a sequestration timescale of at least one year. Assuming a steady state (no accumulation of carbon within the upper ocean), the ANCP obtained with our definition of H can be interpreted as the amount of carbon crossing the depth H horizon and being isolated in the ocean interior for at least one year. We have clarified this background of our study (L106–121), also have discussed the effect of the respiration below the euphotic zone in the result section (L261–271, and L322–335).

In addition to the scientific rationale, there are several practical reasons for using the time-invariant annual maximum MLD as the lower boundary of the upper ocean box in this study. One reason is that by using a time-invariant annual maximum MLD, we can omit the entrainment/detrainment terms associated with temporal changes in the thickness of the box, which is generally difficult to be accurately calculated from monthly data. Additionally, according to our definition, the surface mixed layer remains within the box throughout the year, eliminating the need to include the several order of magnitude larger and thus highly error-prone diffusion within the mixed layer in the budget calculations. This practical consideration has clarified in the method section (L473–485).

4. Comparison to other model-based estimates of export

Previous studies have taken a similar approach as the author's by using a model combined with observations to estimate global export. Devries et al (Glob. Biogeo. Chem., 2017) used a model that fit satellite productivity and observed particle sinking rate, dissolved organic carbon and subsurface O2 and estimated a global export rate of 9 Gt C/yr at the base of euphotic layer. Schlitzer (J. Oceanogr., 2004) used a global circulation model and observed nutrient and O2 fields to estimate a global export rate of 9.6 Gt C/yr at the base of the euphotic layer. The author's need to compare their results to these previous results and offer possible explanations of why their surface O2 based export estimate of 6 GtC/yr is significantly lower than these previous estimates.

Thank you for your comments. As previously stated by Quay (2023) and Palevsky and Doney (2018), the definition of the depth plane for calculating ANCP is crucial. For example, Wang et al. (2023), using an inverse model, showed that only two-thirds of the flux at the bottom of the euphotic layer (73.4 m in their model) reaches 100m due to the strong subsurface respiration. Both studies you referenced use the bottom of the euphotic zone as the lower boundary of the upper ocean box (i.e., depth *H*), which makes direct comparison challenging. One potential reason for the discrepancy in the estimated values may be attributed to the depth definition employed in our study. Although our method does not provide detailed information on the timescales of sequestration, the most recent estimates, derived from an inverse model that combines observations and models (Wang et al. 2023), indicate that the global export flux with sequestration timescales of more than one year is 8.25 ± 0.30 Pg C year^{-1}, which is consistent with our revised estimate, given the range of uncertainties.

We have added a discussion on the factors contributing to the smaller ANCP estimates compared to previous studies in the main text (L261–273, and L393–395) and Supplementary Information (Note S3).

5. Disagreement with observations south of 50° S

Although there is good agreement between the author's results with previous estimates north of 40°S their estimated ANCP is consistently lower than previous estimates south of 40°S. The author's should state possible explanations for this discrepancy.

Thanks to your suggestion (inclusion of a sea level pressure correction in the calculation of the oxygen saturation), the estimation of air-sea oxygen fluxes has been improved, which has significantly reduced the discrepancies with previous studies on ANCP in the Southern Ocean.

Specific issues

Line 425: Did they include spatial and temporal variations in atmospheric pressure when calculating O2 saturation state which is particularly important south of 50°S. The atmospheric pressure term should be included in equation 5.

As you pointed out, the sea level pressure correction in the calculation of saturated oxygen concentration had a substantial impact on the results. In regions with low atmospheric pressure throughout the year, such as the Southern Ocean and high latitudes of the Northern Hemisphere (Fig. S1c), the saturated oxygen concentration is corrected to be lower, resulting in reduced oxygen uptake (red and gray bars in Fig. S6). We have revised the Methods (L489–496)

Line 453: Although they state the use of three different Kz values (4,6 and 8 e-4 m2/s) at the base of the mixed layer to estimate the vertical flux of O2, it wasn't clear how they choose which Kz value to use at a specific location and time.

In the first manuscript, we used three values of the vertical diffusion coefficient at the mixed layer base (which is common globally and time-invariant) to calculate three cases of vertical diffusion flux, and the three values were derived from Cronin et al. (2015).

In the revised manuscript, we used the monthly vertical diffusion coefficients at the mixed layer base from Cronin et al. (2015). As the monthly vertical diffusion coefficients were derived from observations in the Northern Hemisphere, they were applied directly to the grid points of the Northern Hemisphere (north of 20ºN). In the tropical grid points (20°S–20°N), the annual mean value was used throughout the year. Finally, in the Southern Hemisphere (south of 20°S), a six-month shifted monthly coefficient was used.

We have revised the Method section to clarify it (L513–535).

Line 510: They should state the fraction of the grid that was missing due to lack of data.

We have added the fraction of the grid that was not missing to the method section (L584).

Reviewed by Paul Quay

References

Bushinsky, S. M., A. R. Gray, K. S. Johnson, and J. L. Sarmiento, 2017: Oxygen in the Southern Ocean From Argo Floats: Determination of Processes Driving Air-Sea Fluxes. *J. Geophys. Res. Oceans*, **122**, 8661–8682, https://doi.org/10.1002/2017JC012923.

Cronin, M. F., N. A. Pelland, S. R. Emerson, and W. R. Crawford, 2015: Estimating diffusivity from the mixed layer heat and salt balances in the N orth P acific. *J. Geophys. Res. Oceans*, **120**, 7346–7362, https://doi.org/10.1002/2015JC011010.

Palevsky, H. I., and S. C. Doney, 2018: How Choice of Depth Horizon Influences the Estimated Spatial Patterns and Global Magnitude of Ocean Carbon Export Flux. *Geophysical Research Letters*, **45**, 4171–4179, https://doi.org/10.1029/2017GL076498.

Quay, P., 2023: Organic Matter Export Rates and the Pathways of Nutrient Supply in the Ocean. *Global Biogeochemical Cycles*, **37**, e2023GB007855, https://doi.org/10.1029/2023GB007855.

Wang, W.-L., W. Fu, F. A. C. Le Moigne, R. T. Letscher, Y. Liu, J.-M. Tang, and F. W.

Primeau, 2023: Biological carbon pump estimate based on multidecadal hydrographic data. *Nature*, **624**, 579–585, https://doi.org/10.1038/s41586-023-06772-4.

Watson, A. J., U. Schuster, J. D. Shutler, T. Holding, I. G. C. Ashton, P. Landschützer, D. K. Woolf, and L. Goddijn-Murphy, 2020: Revised estimates of ocean-atmosphere CO 2 flux are consistent with ocean carbon inventory. *Nat Commun*, **11**, 4422, https://doi.org/10.1038/s41467-020-18203-3.

Yang, B., S. R. Emerson, and S. M. Bushinsky, 2017: Annual net community production in the subtropical Pacific Ocean from in situ oxygen measurements on profiling floats. *Global Biogeochemical Cycles*, **31**, 728–744, https://doi.org/10.1002/2016GB005545.

——, ——, and M. F. Cronin, 2022: Skin Temperature Correction for Calculations of Air-Sea Oxygen Flux and Annual Net Community Production. *Geophysical Research Letters*, **49**, e2021GL096103, https://doi.org/10.1029/2021GL096103.

Reviewer #3

In Yamaguchi et al. “Global upper ocean oxygen budget and an observational constraint on the biological pump” the authors employ a range of upper ocean oxygen data and modify an upper ocean oxygen mass balance approach to calculate global ANCP. The basic approach is to estimate physical fluxes in the upper ocean oxygen budget and then calculate ANCP as the difference between observed oxygen fluxes and the estimated physical fluxes. They incorporate many of the advances in this approach that have been made over the past decade and calculate uncertainty in their approach using a range of estimates for physical fluxes. This is a natural expansion on past work and represents a promising extension of the approach.

Thank you for taking the time to carefully read our manuscript. In response to your comments, we have made the following revisions:

(1) We recalculated the errors for the terms derived from individual profiles in the dissolved oxygen budget calculation (air-sea flux and vertical diffusion) using a Monte Carlo approach. As the result, the error for the air-sea flux term estimated appropriately larger, while the error for the diffusion term slightly decreased (Fig. 3, and Method).
(2) We introduced a correction for changes in saturated oxygen concentration due to atmospheric pressure in the calculation of the air-sea flux term. This adjustment resulted in reduced oxygen uptake in the Southern Ocean, where sea level pressure is generally low throughout the year, making the latitudinal distribution of ANCP more consistent with previous studies (Fig. S6, and Fig. 4).
(3) According to the changes in the results, we revised related discussions (Section “Global ocean ANCP based on dissolved oxygen budget (L217–)” and “Implications for the global ocean carbon cycle (L354–)”).

For more detailed points and other minor changes, please refer to our responses to individual comments.

My overall main concern with this work is the calculation of uncertainty. I don’t believe the error propagation you use takes into account non-linearities in biases. For example, if different wind speed products give different strength wintertime winds, then the error is not just the variance in the wind speed, but the asymmetry of the wind speed impact on

air-sea fluxes. I think a Monte Carlo is the appropriate way to propagate errors in this situation, not simply variance propagation.
For example: 527-529 – “random measurement errors are considered less problematic if enough observations are utilized for the calculation of means” – true. However, has the work been done to make sure the errors in the oxygen products you use are unbiased? If errors are systematic, rather than random, or have any sort of temporal or spatial component to the bias, this could create serious problems for your analysis that I do not believe are captured in your uncertainty estimate. I would either include the possibility of biased error (not just random) or show that the underlying datasets are not biased (i.e. no systematic differences exist in either Argo vs. Ship or the mapped products vs. underlying observations relative to the results/conclusions you highlight). It seems as though you did not include oxygen measurement uncertainty in your calculation of air-sea fluxes – is that true? At the very least you should include the range of oxygen values that go into your gridded data, but even that does not address the possibility of systematic bias, as mentioned above.

Thank you for your comments. As you and Reviewer #2 pointed out, we recognize that the error estimation in the first manuscript was inadequate. In the revised manuscript, we have changed the error estimation method, primarily following the approach of Bushinsky et al. (2017).

First, we examined the uncertainty and potential bias of the BGC Argo data by comparing these with the GLODAPv2.2022 data. The results of the crossover comparison (data that were matched to within ± 20 days, ± 0.5º latitude, ± 0.5º longitude, and ± 10 dbar) are shown in Fig. S5. Consistent with previous studies (Bushinsky et al. 2017; Maurer et al. 2021), the results indicate no clear offset in the surface oxygen concentration with uncertainty within 1%. Therefore, we assumed a 1% uncertainty for surface dissolved oxygen concentration in the following error estimation.

In the air–sea flux calculation, for each profile, the flux calculations were repeated 1,500 times using randomly selected values. A randomly selected wind value was taken from a normal distribution with an average value of five spatially interpolated daily wind products and a standard deviation of five values. A surface oxygen concentration was also randomly taken from a normal distribution with the observed surface oxygen as the mean and the 1% error of the mean as the standard deviation. The mean of the iterations was taken as the flux value, and its standard deviation was employed as a measure of uncertainty. Flux values and their errors, obtained from 1,500 repeated calculations across ~ 600,000 profiles, were spatially mapped to a 1° (longitude) x 1° (latitude) grid. The errors associated with the globally integrated fluxes are reported as the simple summation of errors mapped.

As the result, The error for the global integrated air-sea oxygen exchange (206 Tmol O_2 year^{-1}) was estimated to be 246 Tmol O_2 year^{-1}. According to these changes, we have revised the result section (Fig. 3 and 4, L196–200, L218–231) and "Uncertainty estimation" in Methods (L588–617).

This is a great effort and a good use of available data. I think this study warrants publication, but my sense is that there may be a good deal more uncertainty in this global approach than is included in the current version. Perhaps that is not the case, but at the very least this should be clearly addressed, otherwise readers may not be convinced that the values shown here are robust and will stand the test of time.

Thank you for carefully reading our initial manuscript. We hope your concerns are addressed by our response as follows.

Seth Bushinsky December 18, 2023

Other important concerns:
Discussion – should include a comparison to Quay 2023, a recent global carbon export study that also used oxygen data (Quay, P. (2023). Organic maber export rates and the pathways of nutrient supply in the ocean. Global Biogeochemical Cycles, 37, e2023GB007855. hbps://doi. org/10.1029/2023GB007855).

Because Quay (2023) does not account for advection and horizontal diffusion fluxes in their oxygen budget calculation, described upper ocean dissolved oxygen mass balance is different from that of this study. In Quay (2023), an oxygen loss from the upper ocean due to the air-sea exchange was estimated ~800 Tmol O_2 year^{-1} and an oxygen loss by vertical diffusion was reported as ~10% of the air-sea flux. These losses were balanced by the net biological oxygen production corresponding to an ANCP of 2.2 ± 0.8 mol C m^{-2} year^{-1}. On the other hand, our results (but recalculated in the area aligned with Quay (2023)) demonstrated that the total ~590 Tmol O_2 year^{-1} of oxygen loss, which is the sum of contributions from air-sea exchanges (~270 Tmol O_2 year^{-1}), diffusion (~190 Tmol O_2 year^{-1}), and advection (~130 Tmol O_2 year^{-1}), should be balanced by the net biological production corresponding an ANCP of 1.4 ± 0.5 mol C m^{-2} year^{-1}. Despite the discrepancy between the two studies, it is noteworthy that both ANCP estimates are consistent in terms of the overall characteristic of the spatial pattern, such as hemispheric and basin contrasts, and the two global estimates, including the range of uncertainty, overlap each other. It is likely because that the contributions from processes not considered in Quay (2023), such as advection and horizontal diffusion, affect the global integrated values of ANCP, yet have a limited impact on determining the basin contrast.

This discussion has been added to Supplementary information (Note S2).

Reproducibility: Code should be shared publicly (for example, using Zenodo). This ties into the incomplete discussion of the mass balance approach I mention below.

We have uploaded all the calculation and plotting codes to Zenodo (https://doi.org/10.5281/zenodo.13148074).

Extended Data Fig. 1 – somethings off here. I recently made a figure of GLODAP O2 profiles per month (see below). Let’s say during the 90s through mid 2010’s GLODAP averaged ~100 profiles per month, that should be over 1000 per year. You figure seems to indicate closer to ~100 per year. How do you handle profiles that appear in multiple databases? Is this because of your order of prioritization? Why would you prioritize WOD over GLODAP when one has had a great deal of attention on accuracy / agreement between data (GLODAP) and the other has many outliers / bad data?

We apologize for the confusion caused by our insufficient explanation in the first manuscript. In the initial manuscript, we prioritized GLODAP over the bottle data from WOD (WOD_OSD) and prioritized the high vertical resolution data from WOD (WOD_CTD) over GLODAP, considering the importance of vertical resolution. Furthermore, our own quality control procedure was applied to exclude profiles with low vertical resolution, which resulted in the use of only approximately 20% of GLODAP data.

In the revised manuscript, we have reviewed the criteria for duplicate checks among the data sources and adopted a protocol that prioritizes GLODAP data over WOD, considering the higher quality of GLODAP data. While the calculation of the time rate change in the dissolved oxygen and vertical diffusion term requires vertical profiles, the flux calculation only requires surface values, resulting in a different number of usable data for each (Figs. S3 and S4). Under this new data utilization protocol, nearly all GLODAP data were used for the air-sea flux calculations (Fig. S3). However, the number of vertical profiles was reduced (Fig. S4) due to the implementation of additional QC to ensure sufficient vertical resolution.

The impact of this change on the results was minimal. These changes in data processing have been reflected in the methods section (L420–436).

Supplemental Information Table 2 – There is a lot of useful information in this table. It would be really helpful to include a scatterplot of your ANCP estimate relative to the local estimates listed here to see how comparable the different approaches are.

The suggestion was excellent. The scatter plot is presented in Figure S7. While there is general agreement in the overall latitudinal trend, as previously observed in Figure 4, with the exception of certain regions (highlighted by the purple box in Figures S7a and b), our estimates tend to be slightly smaller than those of previous local studies. We think that there are two primary reasons for the discrepancies. Firstly, many previous studies estimated ANCP at shallower depths (e.g., seasonally varying mixed layer depth (MLD) and euphotic layer depth (Z_{eu})) than our own (the deeper of the annual maximum MLD and the annual maximum Z_{eu}). Given the definition of the NCP as the net community production (NPP) minus heterotrophic respiration (HR), NCP generally decreases as the depth at which the budget is conducted (referred to as depth H in the manuscript) increases due to the dominance of HR over NPP at depth. When compared with previous studies that employ the identical definition of H, the discrepancy is more modest.

The second reason for this discrepancy is the inclusion of a skin layer temperature correction in the calculation of air-sea oxygen exchange. The incorporation of skin layer cooling into the calculation of surface saturated oxygen concentration results in an increase in the surface saturated oxygen concentration, thereby creating a tendency towards higher oxygen uptake values. According to the budget equation used to calculate ANCP in this study, elevated ocean uptake of oxygen leads a reduction in the net biological oxygen production, that is, ANCP, thereby yielding ANCP estimates that are smaller than those estimated in previous studies. The majority of the previous studies, with the exception of a few (Emerson and Yang 2022; Yang et al. 2022), do not employ the skin layer temperature correction.

This discussion has been incorporated into the main text (L261–273) and the supplementary information (Note S3).

Line 102 – Should state a bit more detail here or later in the upper ocean mass balance section about how the seasonal cycles are initially derived. I think you grid the observations and then do the mass balance calculation, but this is not clear and is an important step.

We have revised the relevant sections in the main text (L99–102) and Methods (L465–471) accordingly.

Impact of setting Southern Ocean physical fluxes to zero – more explanation of why this calculation is done and what you are showing with this is needed. It comes out of nowhere and setting physical fluxes to zero in the SO seems like an odd null hypothesis.

In the first manuscript, the associated discussions were based on the assumption of zero physical flux in the Southern Ocean in order to assess the impact of the discrepancy with previous studies in the Southern Ocean on the global ANCP estimate. In the revised manuscript, the air–sea oxygen flux calculations have been revised to incorporate a comment made by Reviewer #2, which recommends the explicit inclusion of sea level pressure in the air-sea flux calculation (L489–496, and Fig. S6). The change has reduced the discrepancy with previous studies, particularly in the Southern Ocean (Fig. 4). Consequently, it was determined that the discussion was no longer necessary and the relevant discussion has been removed from the manuscript.

Minor comments:

54 – “an understanding of the processes” – I think this is supposed to be one of the clauses following “lack of”... but it is unclear as written and makes it sounds like there is an understanding of the processes. Please clarify the intent of this sentence.

We have revised the text as follows (L49–52): “Global observational quantification of the carbon fixed and exported to the deep ocean through the biological pump remains insufficiently constrained due to a lack of direct observations that can be mapped globally and a lack of understanding of the processes that can be modeled.”.

Paragraph starting line 52 – I don’t this that comparison of biological carbon export uncertainty to air-sea co2 flux uncertainty is fair, or at least not pinning the difference on a difference in in situ observational data. One calculation involves mapping a single quantity and calculating a single physical flux due to a difference in partial pressure, the other has requires much more than concentration gradients to estimate.

As you pointed out, a direct comparison of two fluxes with disparate data requirements and differing difficulties in data collection may be unfair. Nevertheless, we wish to emphasize that both fluxes are equally important for quantitatively describing the global carbon cycle and that estimates of carbon fluxes within the ocean interior are relatively less constrained. We have revised the text as follows to clarify the reasons beyond just the lack of observational data (L56–60): “This weak observational constraint stands in contrast to the relatively well-constrained global estimate of air–sea CO_2 exchange for the recent decade (2.8 ± 0.4 Pg C year^{-1}, mean ± SD), despite both being essential components of the global carbon cycle. This is partly because of the difference in the number of observations required to quantify them due to the different complexities among their underlying processes.”

70 – lifetime is more like 5-7 years

We have revised it (L70).

88-89, 157-19 – Note that Bushinsky and Emerson 2015 and many subsequent studies included lateral and vertical physical processes, so these statements are not correct as written. Repeatedly mentioned as “overlooked” throughout the text, though this is not correct.

We agree. We have revised the several relevant parts (L87–90, L175–176, and L233–238).

119 – cold temperature and high solubility on its own doesn’t induce undersaturation. The rate of cooling must exceed the replenishment of oxygen from air-sea exchange.

We have revised the text as follows (L134–137):” The strong air–sea oxygen gradient in the cold season is a consequence of undersaturated sea surface oxygen, which results from the rapid cooling and continuous supply of relatively low oxygen water from below due to the surface cooling and associated convection.”

Fig.4 - confusing legend in (b) – are the blue, red, and yellow lines all showing the same quantity (ANCP/ANPP) using different estimates of ANCP?

We apologize for the confusion. The differing terminology used in the initial manuscript is due to the fact that these three quantities have different equation forms for the parameterization and input variables to the parameterization, although they can all be referred to as "export ratio." Laws et al. (2011) use a parameterization of the export ratio (referred to as "ef-ratio" in the paper) based on sea surface temperature (SST) and NPP, whereas Li and Cassar (2016) use a parameterization of NCP based on SST and NPP. Using these parameterizations, export ratios (ANCP/ANPP) calculated from common SST and NPP are shown in Figure 6. We have added this explanation in the figure caption.

Line 325 – “trends of both” – both what? ANCP estimates, or ANCP and NPP?

We have revised the text as follows (L357–360): “The latitudinal trends of both ANCP and ANPP exhibit a decrease from the equator northward followed by an increase, showing a coherent pattern in the NH (Figs. 4f and 6).”

References

Bushinsky, S. M., A. R. Gray, K. S. Johnson, and J. L. Sarmiento, 2017: Oxygen in the Southern Ocean From Argo Floats: Determination of Processes Driving Air-Sea Fluxes. *J. Geophys. Res. Oceans,* **122**, 8661–8682, https://doi.org/10.1002/2017JC012923.

Emerson, S., and B. Yang, 2022: The Ocean's Biological Pump: In Situ Oxygen Measurements in the Subtropical Oceans. *Geophysical Research Letters*, **49**, e2022GL099834, https://doi.org/10.1029/2022GL099834.

Laws, E. A., E. D'Sa, and P. Naik, 2011: Simple equations to estimate ratios of new or export production to total production from satellite-derived estimates of sea surface temperature and primary production. *Limnology and Oceanography: Methods,* **9**, 593–601, https://doi.org/10.4319/lom.2011.9.593.

Li, Z., and N. Cassar, 2016: Satellite estimates of net community production based on O2/Ar observations and comparison to other estimates. *Global Biogeochemical Cycles*, **30**, 735–752, https://doi.org/10.1002/2015GB005314.

Maurer, T. L., J. N. Plant, and K. S. Johnson, 2021: Delayed-Mode Quality Control of Oxygen, Nitrate, and pH Data on SOCCOM Biogeochemical Profiling Floats. *Front. Mar. Sci.*, **8**, https://doi.org/10.3389/fmars.2021.683207.

Quay, P., 2023: Organic Matter Export Rates and the Pathways of Nutrient Supply in the Ocean. *Global Biogeochemical Cycles*, **37**, e2023GB007855, https://doi.org/10.1029/2023GB007855.

Yang, B., S. R. Emerson, and M. F. Cronin, 2022: Skin Temperature Correction for Calculations of Air-Sea Oxygen Flux and Annual Net Community Production. *Geophysical Research Letters,* **49**, e2021GL096103, https://doi.org/10.1029/2021GL096103.

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

The authors have generally done a nice job of responding to the reviews, particularly to comments about uncertainty and statistics. It is a very nice piece of work and has global significance. However, I will make one very critical comment that seem easily fixable. That regards my comment with respect to their decision to calculate Annual Net Community Production by integrating to the maximum winter mixed layer depth. That decision leads to a systematic bias low in their results because the mass balance they use includes significant amounts of respiration below the euphotic zone. They acknowledge that "our ANCPs are on the smaller side" on lines 270/273 due to the "greater influence of subsurface respiration" and this is clearly shown in Figure S7a where they compare their results to other studies. It shows a bias low of 1/3, which I believe is a new addition to the supplementary material.

Annual net community production (ANCP) is defined as the difference between net primary production (NPP) and respiration by heterotrophs (animals and bacteria) in the upper ocean. ANCP is also equated to the amount of carbon that is exported from the upper ocean over an annual cycle. For example, as stated in Emerson and Hedges (Chemical Oceanography and the Marine Carbon Cycle, 2008, pg. 30), "the mean organic carbon flux out of the euphotic zone should equal the net community production of organic carbon". Note that this clearly defines the upper ocean in the determination of NCP as the euphotic zone. Unfortunately there are a number of studies of ANCP, mostly based on tracer budgets, that have used the maximum depth of the mixed layer to define ANCP. That happens for reasons explained in this manuscript (one doesn't have to calculate entrainment and detrainment). These studies primarily occur at places such as Ocean Station Papa or the Hawaii Ocean Time-series site, where the maximum mixed layer depth is not particularly deep and not strikingly different than the euphotic zone. So the biases aren't large in these other studies and have largely eluded comment. However, in the case of the global analysis in this manuscript, much of the ocean has quite deep mixed layers. Figure 3 shows everything poleward of 40 degrees as deeper than 200 m and well below the euphotic zone. No wonder their global ANCP seems low.

Does this mean that the results of the paper are invalid? No, it's a really nice piece of work. The authors could fix it simply by focusing on their statement that their results are "the equivalent amount of carbon, in either particulate or dissolved form, escapes from the predominant upper-ocean seasonal cycle of physical environmental conditions (e.g., mixing and restratification) and becomes sequestered into the ocean interior on longer timescales than, at shortest, one year.” (lines 202/204). That is a critical value for our understanding of the global carbon cycle. Rather than calling the results ANCP, maybe use something like Annual Carbon Sequestration? They could get credit for coining the term. They could clearly note that their method is highly related to ANCP and yields a lower limit. But they should avoid trying to equate it to ANCP.

Thank you for the detailed explanation. And, thank you also for your specific solution suggestions. We understand your concern.

What we have obtained from the upper ocean oxygen budget calculation in this study is simply the "annual net biological oxygen production" integrated from the surface to the annual maximum mixed layer depth. By multiplying the oxygen production by the assumed oxygen carbon ratio, we can then estimate the "annual net organic carbon production" in the layer. In order to interpret this value estimated for each grid point as "annual carbon sequestration", another strong assumption is needed: that all organic carbon produced in the layer is transported to the layer below.

In terms of time, this assumption is equivalent to the condition that there is no net accumulation of organic carbon within the layer over the course of a year, and is therefore a relatively easy assumption to make over a large area of the present-day ocean. On the other hand, the assumption does not necessarily hold spatially. For example, where horizontal current velocities are high, there is likely to be an annual net horizontal transport of organic carbon. In other words, Figure 4a cannot be interpreted directly as "annual carbon sequestration" in each grid.

Because of this problem, we avoided referring to the estimated "annual net organic carbon production" for each grid point as "annual carbon sequestration" in the manuscript. We interpreted as "annual carbon sequestration" only those globally integrated values where horizontal transport is completely cancelled out, or averaged over spatial scales where the effects of horizontal transport would be negligible. Therefore, we are reluctant to replace all "ANCP" in the manuscript with the proposed "annual carbon sequestration". On the other hand, we are also reluctant to define new terms in this manuscript for fear of further confusing the reader.

As a compromise, we add an additional note in the Introduction (L122) and a subscript H to ANCP throughout the manuscript to keep this caveat in mind for the readers throughout the manuscript.

Reviewer #3 (Remarks to the Author):
Thank you to the authors for their comprehensive responses to the reviewer suggestions. I think the current version of the manuscript has sufficiently responded to all comments and is ready for publication. I would also like to add that I think their use of the seasonal maximum mixed layer depth for their ANCP calculation makes sense and is the correct approach for this type of work, given their research questions.

Seth Bushinsky

Thank you for carefully reading the manuscript and for your helpful comments.