Dear Reviewer 1,

We are grateful for your thoughtful and thorough evaluation of our manuscript. Below, we respond to each of your comments and concerns, identifying how we intend to address them in a revised version of the manuscript. Our responses are indicated with a blue font. We have endeavoured to address all of your comments as you recommended. Overall, we believe that our manuscript will be improved by addressing your comments and we thank you again for your time reviewing this document.

Please note the accompanying updated manuscript in which our revisions are highlighted in yellow. In our responses below, line numbers refer to those of the revised manuscript.

Sincerely,
R. Izett & co-authors.

Specific comments:

Line 44: choose a different word than 'sinks' as not all export is a sinking flux

Thank you for the suggestion. We changed the sentence, beginning on revised manuscript L46, from

"*When measured over sufficiently large temporal and spatial scales, NCP quantifies the amount of photosynthetically produced organic matter that sinks from the upper ocean (Laws 1991)*"

to

"*When measured over sufficiently large temporal and spatial scales, NCP quantifies the amount of photosynthetically produced organic matter that is removed from the upper ocean (Laws 1991)*".

Equation 1,2,3: it seems counterintuitive to have CR added to GPP in equations 1 & 2 rather than subtracted. I think the reason it is shown this way is because CR is assumed to have a negative value. However, on page 8 it is stated that the first term on the right of equation 5.1 = GPP and the second term on the right = CR, making the relationship GPP – CR, which is inconsistent with equations 1 & 2. It also seems nonintuitive in equations 1 and 3 to have a + sign in front of the last term on the right for equation 1 and a minus sign for the last term of equation 3. Should these both be '+/-' since they represent source and sink terms?

Equation 7.1: should '+/-' be used in front of the final term on the right rather than '-' ?

Thank you for looking closely at the equations. In response to your suggestions, we changed all respiration terms (AR, HR, CR) to have positive notation throughout the manuscript, such that NPP = GPP-AR and NCP = GPP-CR. To clarify this notation, we changed the last sentence of the second introductory paragraph (beginning L46) from

"*All PP fractions are often expressed as volumetric equivalents of organic carbon or O2 production (e.g., mol C or O2 m-3 d-1), such that respiration has negative values*"

to

"*GPP, NPP and NCP are often expressed as volumetric equivalents of organic carbon or O2 production (e.g., mol C or O2 m-3 d-1), and respiration terms are expressed in terms of* *organic C or O2 consumption. Accordingly, in a closed system, GPP, NPP and CR can only have positive values, while NCP may assume positive or negative quantities.*"

We also changed the "- other source/sinks" term in the equations to "± other sources and sinks"

As a result of these changes, equations 1-3 and 7 become:

$$d[T(t,z)]/dt = GPP(t,z) - CR(t,z) \pm \text{other sources/sinks}(t,z) \qquad (1)$$

$$d[T(t,z)]/dt \approx GPP(t,z) - CR(t,z) \qquad (2)$$

$$NCP(t,z) = d[T(t,z)]/dt \pm \text{other sources/sinks}(t,z) \qquad (3)$$

$$NCP(t,z) = (h_{i+1} - h_i)[T(t_1,z)] - [T(t_0,z)]\ t_1 - t_0 \pm \Sigma F(t,z) \qquad (7)$$

Line 160 – 164: Here it is stated that POC is estimated from published relationships (Loisel et al, Cetinic et al, Graff et al.). I would suggest explicitly giving these relationships in a table in the appendix. The Graff et al. paper, for example, is primarily focused on estimating phytoplankton carbon from bbp and the POC relationship is a secondary result. Explicitly providing the equations used will prevent any confusion.

Thank you for the suggestion. We included the following supplementary table, and associated references, in the appendix of the manuscript.

Table A4. A comparison of selected c_p- and b_{bp}-to-POC algorithms. Resulting POC units are mg m^{-3}. Units of c_p and b_{bp} are both in m^{-1}, and the wavelength of the c_p and b_b measurements is indicated with a subscripted number (e.g., $c_{p,660}$ indicates measurements at 660 nm). This table is not a complete list; the equations were selected to illustrate variability in POC relationships.

POC Equation	**Region**	**Reference**
$POC = 367\ c_{p,660} + 31.2$	N. Atlantic	Marra et al. (1995)
$POC = 391\ c_{p,660} - 5.8$	N. Atlantic	Cetinić et al. (2012)
$POC = 35422\ b_{bp,700} - 14.4$	N. Atlantic	Cetinić et al. (2012)
$POC = 48811\ b_{bp,470} - 24$	N. and S. Atlantic, Equatorial Pacific	Graff et al. (2015)
$POC = 841\ b_{bp,532}^{0.395}$	N. and S. Atlantic	Balch et al. (2010)
$POC = 39418\ b_{bp,470} - 13$	S. Atlantic; Southern Ocean	Thomalla et al. (2017)
$POC = 501.81\ c_{p,660} + 5.33$	Equatorial Pacific	Claustre et al. (1999)
$POC = 585.2\ c_{p,660} + 7.6$	Equatorial Pacific	Behrenfeld and Boss (2006)
$POC = 661.9\ c_{p,660} - 2.168$	Pacific and Atlantic (incl. upwelling)	Stramski et al. (2008)
$POC = 71002\ b_{bp,555} - 5.5$	Pacific and Atlantic (incl. upwelling)	Stramski et al. (2008)
$POC = 458.3\ c_{p,660} + 10.713$	Pacific and Atlantic (excl. upwelling)	Stramski et al. (2008)
$POC = 53932.4\ b_{bp,555} - 5.049$	Pacific and Atlantic (excl. upwelling)	Stramski et al. (2008)
$POC = 574\ c_{p,555} - 7.4$	Mediterranean	Oubelkheir et al. (2005)
$POC = 404\ c_{p,660} + 29.25$	Mediterranean	Loisel et al. (2011)
$POC = 37550\ b_{bp,555} + 1.3$	Mediterranean	Loisel et al. (2011)
$POC = 31200\ b_{bp,700} + 3.04$	Southern Ocean	Johnson et al. (2017)
$POC = 977760\ b_{bp,770}^{1.166}$	Southern Ocean	Johnson et al. (2017)
$POC = 17069\ b_{bp,555}^{0.859}$	Antarctic Polar Frontal Zone	Stramski et al. (1999)
$POC = 476935.8\ b_{bp,555}^{1.277}$	Ross Sea	Stramski et al. (1999)
$POC = 381\ c_{p,660} + 9.4$	Global Ocean	Gardner et al. (2006)

Line 493: Check the wording of the sentence beginning ‘Our calculation…’, something is wrong here

Thanks for pointing this out. We changed the sentence, beginning on L521, from

“*Our calculations, we extend the work of Johnson and Bif (2021) and Stoer and Fennel (2022)*”

to

“*Our calculations extend the work of Johnson and Bif (2021) and Stoer and Fennel (2022)*”

General: When calculations of production are made where nighttime changes in a given tracer are assumed to be applicable to daytime rates, what error might be introduced because of impacts of diel vertical migrators?

In section 4.1, we discuss potential uncertainty in diurnal-cycle based GPP calculations. In the second paragraph (beginning L602), we discuss uncertainty resulting from non-photosynthetic processes that vary diurnally, such as air-sea gas flux, grazing, and sinking. In response to your feedback, we modified this paragraph to specifically include grazing and diel vertical migration. A revised paragraph is as follows (new text is underlined, with additional re-phrasing throughout to improve clarity):

Diurnal cycle GPP methods are based on the presumption that day-night variations in photosynthesis are the primary driver of diurnal variations in O_2 or POC concentrations in the upper ocean. Other than accounting for potential diurnal solubility impacts on O_2 (through expressing O_2 as its concentration anomaly, ΔO_2) no attempts have been made to reconcile for additional diurnal variations in float O_2 or POC observations that are not caused by photosynthesis. For O_2, these include potential impacts due to air-sea exchange or vertical mixing, and for POC, sinking, diel vertical migration and grazing, or PER. Yet, these processes vary throughout the day, and the extent to which they do depends on the season and region. Diurnal variability in solar heating and wind forcing influence mixed layer dynamics on hourly, or longer, timescales, with impacts on air-sea gas exchange (Briggs et al., 2018; Barone et al., 2019) and near-surface vertical mixing (Price et al., 1986). Moreover, particle sinking, grazing, and DOC production, have been implicated as a mechanism for decoupling O_2- and POC-based PP estimates, particularly in high-productivity (e.g., diatom-dominated) regions (e.g., Rosengard et al., 2020). For example, regions of high POC sinking rates, grazing or PER will decouple O_2 and POC concentrations, leading to observations of high-O_2 and low-POC in upper ocean waters, with implications for resulting GPP and CR estimates (White et al., 2017; Rosengard et al., 2020; Briggs et al., 2018). Similarly, day-night variations in grazing, resulting from diel vertical migrations, could amplify the nighttime decline in POC, thereby artificially inflating nighttime respiration estimates, and decoupling O_2- and POC-based GPP calculations. Independently or in combination, these non-photosynthesis diurnal processes likely imprint on the daily signals detected by BGC-Argo floats, whether by single assets or the composite of the array, and therefore constitute a source of uncertainty to the resulting GPP estimates.

Line 760: Since the previous statements include assessments of satellites, it is not clear what is implied by stating that float data are ‘publicly available’ since satellite data are also publicly available.

Our intention in this statement was to compare the availability of float data versus ship/bottle data. While recent efforts towards FAIR data principles have improved the availability of ship/bottle data, they remain less accessible (e.g., spread over multiple, disconnected repositories) and not standardized (e.g., bottle/ship PP datasets are often published individually with a single paper/project, and therefore follow no archiving or metadata guidelines). We clarified these points in the opening sentences of the conclusions, as in the following revised paragraph (new text is underlined).

The BGC-Argo fleet offers global observations of real-time ocean biogeochemistry, enabling widespread PP measurements that are independent of, yet complementary to satellite and ship-based approaches. However, compared with PP methods that rely on traditional sampling infrastructure, float-based methods confer significant advantages in detecting PP. Float-based methods, for example, provide simultaneous horizontal, vertical, and temporal PP coverage, presenting the opportunity to fill key gaps in the existing PP data record (Fig. 1). Moreover, while recent efforts towards FAIR data principles (Tanhua et al., 2019) have improved the availability of ship and bottle data, resulting PP datasets remain generally inaccessible (e.g., spread over disconnected repositories) and non-standardized (e.g., datasets are often published individually with a single paper/project, and therefore follow no archiving or metadata guidelines). Float data, in contrast, are generally made available within 24 hours of collection, are publicly available and are archived following agreed-upon guidelines (Bittig et al., 2019), enabling cost-effective, open-source PP calculations that can be independently verified and applied by the entire science community, including those without the resources to perform traditional PP methods. Lastly, float-based methods facilitate enhanced detection of the biological response to unpredictable or episodic events like wildfires, volcanic eruptions, or bloom periods, which often cannot be sufficiently characterized using traditional in-situ datasets (Tang et al., 2021).

Bittig, H. C., Maurer, T. L., Plant, J. N., Schmechtig, C., Wong, A. P. S., Claustre, H., Trull, T. W., Udaya Bhaskar, T. V. S., Boss, E., Dall'Olmo, G., Organelli, E., Poteau, A., Johnson, K. S., Hanstein, C., Leymarie, E., Le Reste, S., Riser, S. C., Rupan, A. R., Taillandier, V., Thierry, V., and Xing, X.: A BGC-Argo Guide: Planning, Deployment, Data Handling and Usage, Frontiers in Marine Science, 6, https://doi.org/10.3389/fmars.2019.00502, 2019.

Tanhua, T., Pouliquen, S., Hausman, J., O'Brien, K., Bricher, P., de Bruin, T., Buck, J. J. H., Burger, E. F., Carval, T., Casey, K. S., Diggs, S., Giorgetti, A., Glaves, H., Harscoat, V., Kinkade, D., Muelbert, J. H., Novellino, A., Pfeil, B., Pulsifer, P. L., Van de Putte, A., Robinson, E., Schaap, D., Smirnov, A., Smith, N., Snowden, D., Spears, T., Stall, S., Tacoma, M., Thijsse, P., Tronstad, S., Vandenberghe, T., Wengren, M., Wyborn, L., and Zhao, Z.: Ocean FAIR Data Services, Frontiers in Marine Science, 6, https://doi.org/10.3389/fmars.2019.00440, 2019.

Line 775: I'm not sure I would advocate using BioArgo production products to train satellite algorithms as my guess is that there is more error/uncertainty in the former than in the latter. I do not see evidence in the current manuscript to conclusively demonstrate otherwise.

We believe that as float-based PP methods mature - and their uncertainties become reduced or better constrained - it will become feasible to train and validate satellite algorithms using float PP data. The resulting algorithms would constitute an entirely independent method to quantifying PP that does not rely on ship-based observations. Similarly, given the on-going expansion of the BGC-Argo array and the continued generation of significant amounts of biogeochemical data, the algorithms can be continually re-trained and evaluated using new methods and datasets. We incorporated these comments in the final paragraph of the conclusions, beginning on L816, as follows (new text is underlined):

Ultimately, continued efforts towards expanding and refining float-based PP datasets will reduce uncertainties in the present methods, yielding widespread, in-situ PP estimates in most ocean basins. As uncertainties are further constrained, the resulting estimates will convey significant tangential benefits, like the ability to improve numerical model predictions through data assimilation (e.g., Wang et al., 2020a) and to train and/or validate satellite PP algorithms, as has been done previously using ship data (e.g., Li and Cassar, 2016; Huang et al., 2021). Given the on-going expansion of the BGC-Argo array and the continued generation of significant amounts of biogeochemical data, the resulting products can be continually re-trained and evaluated using new methods and datasets. Achieving these milestones will enable unprecedented, in situ classification of the response and variability of marine PP to various environmental perturbations over a range of space and time scales.

Grammar

Line 223: add 'relationship' after photosynthesis-versus-irradiance

Line 229: Define OSP on first use

Line 348: replace 'are' with 'is'

Line 445: add 'is' after 'values'

Figure 5: define 'Y17', 'H22' and 'H20'

Line 532: replace 'of' with 'our'

Line 569: add 'between' after 'observed.'

Line 690: Delete 'And' at the beginning of the sentence and just begin with 'To'

Line 771: add 'be' after 'can'

Thank you for identifying these mistakes. We made all of these changes as you have suggested, including defining Y17, H22 and H20 in the figure 5 caption.

Dear Reviewer 2,

We are grateful for your thoughtful and thorough evaluation of our manuscript. Below, we respond to each of your comments and concerns, identifying how we intend to address them in a revised version of the manuscript. Our responses are indicated with a blue font. We have endeavoured to address all of your comments following your recommendations. In some cases, however, we feel the recommended changes are unnecessary, and explain why we think so. We believe that our manuscript will be improved by addressing your comments and we thank you again for your time reviewing this document.

Please note the accompanying updated manuscript in which our revisions are highlighted in yellow. In our responses below, line numbers refer to those of the revised manuscript.

Sincerely,
R. Izett & co-authors.

General:

This manuscript aims to provide an overview of available estimates of gross primary production (GPP) and net community production (NCP) obtained from the analysis of BGC-Argo float data. The manuscript starts with a detailed description of the assumptions driving GPP and NCP measurements based on float data, intertwined with a description of some of the existing studies reporting such productivity estimates for different ocean regions (and globally). In the second part of this manuscript (section 3 onwards), the authors further review available NCP estimates for OSP, and conduct their own novel analysis to derive global GPP estimates. Overall, I believe there is a lot of useful information compiled in this work, but the different sections of the manuscript feel disconnected from one another. Moreover, regarding the review of NCP estimates, the authors omit several studies that have inferred respiration rates in the mesopelagic layer from float data and have been used to obtain NCP estimates in the Southern Ocean. My initial recommendation is to divide the present manuscript into two separate works, dealing with GPP and NCP, separately. At this stage, I find the work related on GPP to be potentially more robust than that on NCP.

Thank you for your assessment of our manuscript. We appreciate and value your feedback; however, we elected to not divide the manuscript into separate contributions for GPP and NCP. Our reasons are the following:

- GPP and NCP are important metrics, which have both benefitted from recent efforts to quantify them using float observations. While the methods to estimate both metrics are different, there are some consistencies, such as the requirement to interpret upper ocean biogeochemical budgets (albeit over different timescales) and in the variables used (O2, bbp, in particular). For these reasons, we feel it makes sense to streamline descriptions of their respective methods, as we have done in the opening paragraphs of section 2.
- Despite the recent attention to developing GPP and NCP methods, the approaches are still new, and not widely known or consistently applied. While promising, both sets of methods require continued attention to limit their uncertainties, and promote their

uptake by the community. We feel that presenting GPP and NCP methods together is most appropriate for summarizing these shared challenges and efforts to address them.

- NCP and GPP are rarely compared. However, our manuscript offers an early opportunity to do this using float data (as in Fig. 10). While our GPP vs NCP analysis is fairly simple, we feel that alongside the other GPP and NCP examples in the manuscript it offers an example of how the methods can be used as powerful tools - independently *and* together - to resolve PP over a range of scales.

Overall, our goal with this manuscript is to provide a single, comprehensive, and accessible reference on emerging float-based GPP and NCP methods. We are targeting a broad readership, including researchers who do not normally perform PP calculations, with the intention of summarizing the current state of GPP and NCP methods, and helping to familiarize the community-at-large with these new tools. Thus, we elected to describe both methods in a singular resource that serves as both an overview to people unfamiliar with float-based PP methods, and as a resource to those familiar with the methods who may wish to understand, at a higher level, the current main benefits and challenges. Ultimately, we hope that this resource facilitates broader uptake of the methods, as singular or combined tools, and promotes their continued development.

We clarified this intention by modifying the following sections of text (revised text underlined).

We added the following text to the end of the abstract (beginning L21): *This paper aims to facilitate broader uptake of float GPP and NCP methods, as singular or combined tools, by the oceanographic community and to promote their continued development..*

We added the following text to the end of the final introduction paragraph (beginning L100): *This paper is intended as a resource for a broad readership — including researchers who do not normally perform PP calculations — that summarizes the current state of GPP and NCP methods and helps to familiarize the community-at-large with the current benefits, challenges and application of these new tools.*

We modified the second paragraph of the conclusions as follows (beginning L808): *As float-based techniques mature, the BGC-Argo fleet can be used to extend our current understanding of the marine GPP, NPP, NCP, and C-export, particularly at scales that have so far only been achieved through satellite-based algorithms (e.g., Behrenfeld and Falkowski, 1997; Laws et al., 2011). For example, by compiling the data discussed and derived in this paper, we can calculate independent, global estimates of the carbon export ratio (equivalent to ANCP divided by NPP, where NPP is derived from float-GOP; Figure 10). Notwithstanding the regional and temporal biases in current float-based PP estimates, these C-export ratio estimates are consistent with the commonly used satellite models of Laws et al. (2011) and Henson et al. (2012). Simultaneous estimates of GPP, NCP, and C-export are rarely made, let alone comparisons between them. Thus, the export ratio we derived here could be an important tool for improving our understanding of the ocean carbon cycle.*

Moving forward, the extent to which float-based PP calculations can be applied will depend, to a large degree, on the availability of float data (sect. 4.1.1), and our capacity to better constrain key sources of uncertainty in biogeochemical budget interpretations (sect. 4.1 and 4.2). Indeed, to increase the availability of float-based PP data, expansion of the Argo fleet should be prioritized, particularly in under-sampled ocean regions. Floats will need to be deployed with sampling intervals set to 5.2 or 10.2 days (rather than 5.0 or 10.0 days) to properly detect diurnal variability. Finally, fully exploiting floats for PP measurements will rely on the open availability of PP datasets, including processed data and relevant software.

Please also see below our responses to your specific comments regarding the omitted Southern Ocean literature. In brief, our intention in this manuscript was to highlight upper ocean (e.g., euphotic zone) processes. However, following your recommendations, we have included references to, and brief descriptions of, the sources that you provide.

Specific comments:

The introductory section is quite complete and provides a good description of the main productivity estimates that one can find in the literature (PP, GPP, NPP, NCP, etc.). For the most part, I like Figure 1, but I do not quite understand the second set of lines below the grey line indicating "autotrophic". I would recommend simplifying panel (a) by removing the last three lines.

Thank you. In Figure 1, our intention was to illustrate a case where CR > GPP, giving rise to net heterotrophic (negative NCP) conditions. The lower part of the figure supports text in the secondary instruction paragraph (starting L46) describing units and notation for the PP fractions. Note that we changed the notation around the respiration terms (see comment below), such that AR, HR and CR have positive notation, and NPP = GPP-AR and NCP = GPP-CR. Accordingly, the text beginning on L46 becomes "*GPP, NPP and NCP are often expressed as volumetric equivalents of organic carbon or O2 production (e.g., mol C or O2 m-3 d-1), and the respiration terms are expressed in terms of organic C or O2 consumption. Accordingly, GPP, NPP and CR can only have positive values, while NCP may assume positive or negative quantities*", and equations 1-3 become:

$$d[T(t,z)]/dt = GPP(t,z) - CR(t,z) \pm \text{other sources/sinks}(t,z) \qquad (1)$$

$$d[T(t,z)]/dt \approx GPP(t,z) - CR(z) \qquad (2)$$

$$NCP(t,z) = d[T(t,z)]/dt \pm \text{other sources/sinks}(t,z) \qquad (3)$$

While we appreciate your feedback, for the reasons described above, we will maintain the lower part of Figure 1, as is. We do, however, clarify in the figure caption what the second set of lines represent. A revised figure caption is as follows, with new text underlined.

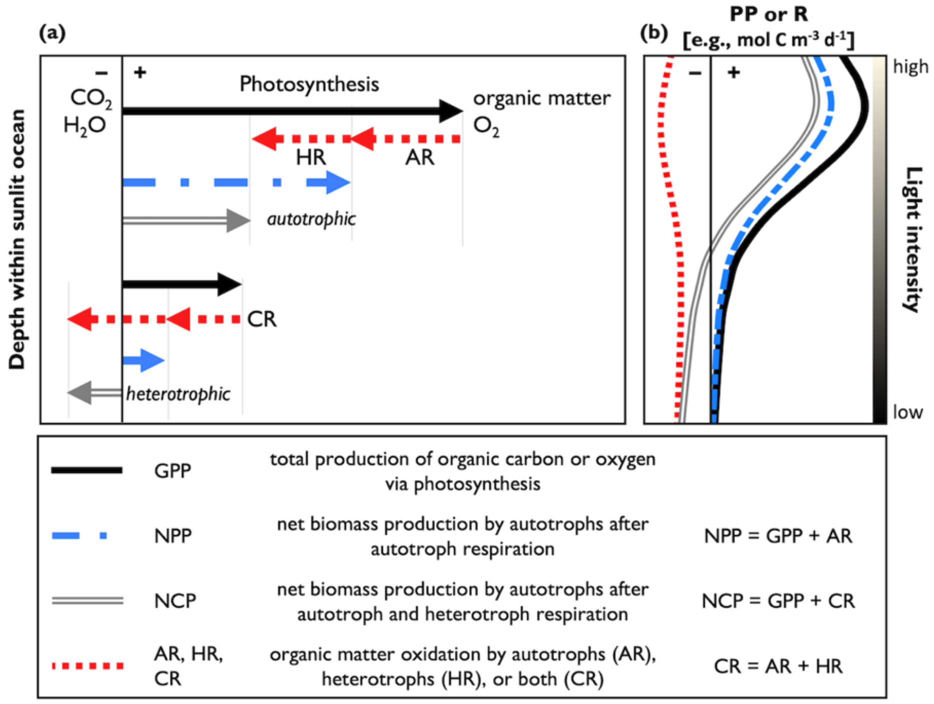


Figure 1. A conceptual schematic of PP definitions. Panel (a) shows simplified reaction equations of organic matter production and respiration. The upper part of the figure represents a region of net autotrophic conditions (NCP > 0), while the lower part represents a region of net heterotrophic conditions (NCP < 0). Note that net heterotrophic conditions do not necessarily always occur deeper in the water column than net autotrophy. Panel (b) represents idealized PP and CR profiles, where PP declines with depth due to the light dependency of photosynthesis. The vertical axis represents water column depth, and the thin black line divides positive and negative rates.

Figure 2d does not make sense to me given the large deployment of floats and profiles made available through the SOCCOM program. Based on this panel, it seems as if the Southern Ocean is one of the least sampled regions in terms of BGC-Argo profiles, which is not the case. I have attached below a map from the GO-BGC website showing that the Southern Ocean is the region with the largest quantity of floats (and thus BGC profiles) (https://www.go-bgc.org/arraystatus#locations). Is Figure 2d perhaps yielding a misleading picture based on the way the data was binned? Furthermore, Figure 2d seems to be inconsistent with Figure 6a, where the largest number of profiles is indeed observed in the Southern Ocean.

Thanks for this. We agree that the original Fig. 2d was somewhat misleading. As you suggest, this is mostly due to the way that the data were binned; and also due to the colour scale on the original figure. We have re-made Fig. 2d, binning the float profiles to 10x10 degree bins and normalizing the results by the surface area in each grid cell. We also adjusted the colour axis for that panel. Now, we believe that the large number of profiles from the Southern Ocean is better represented, as in Fig. 6a. Moreover, our heat map in Fig. 2d is consistent with a recent snapshot of BGC-Argo profile data obtained from the network status map (see below figure).

Revised Figure 2 and caption (revised text underlined):

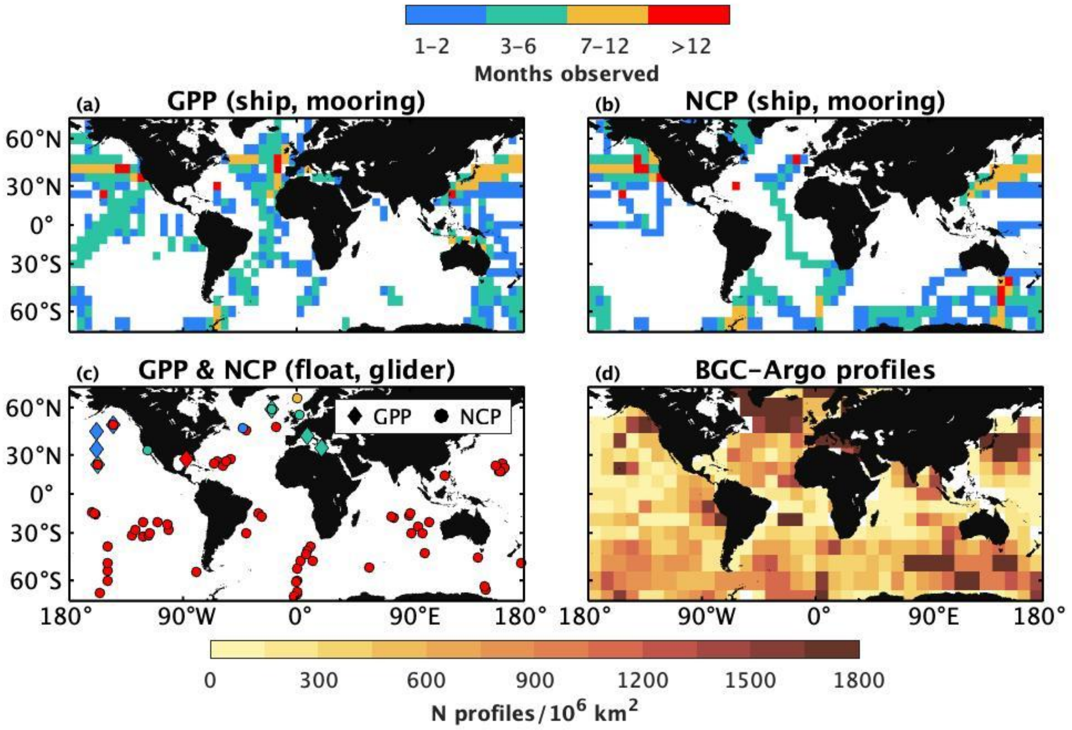
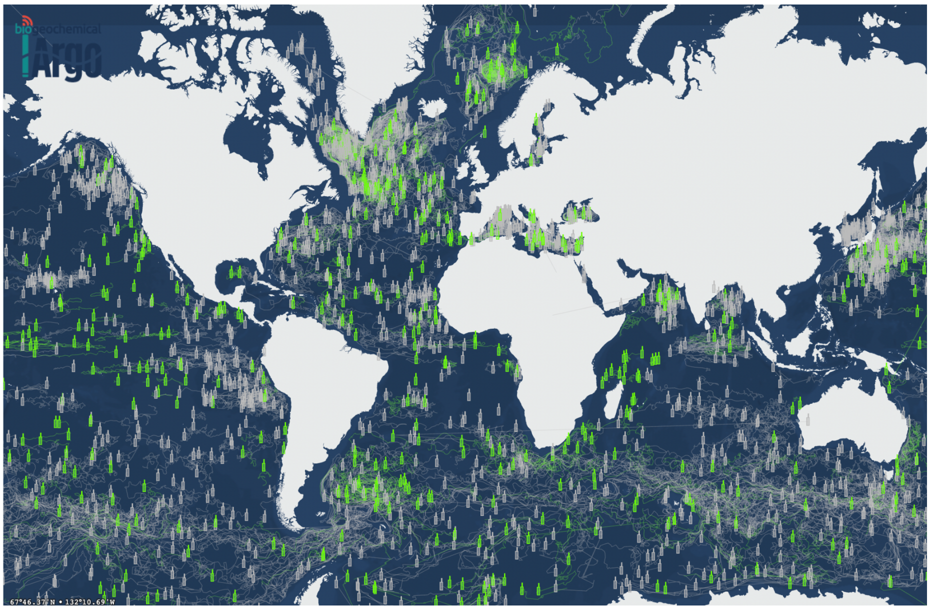


Figure 2. Coverage of GPP and NCP datasets, and BGC-Argo profiles. The upper row represents archived GPP and NCP data obtained from ships or moorings, while panel (c) shows the locations and durations of float- or glider-based GPP and NCP studies. Panel (d) shows a heatmap of the distribution of BGC-Argo profiles collected from 2010 through 2022. Data in panels (a) and (b) were binned to a five-by-five-degree grid. Data in panel (d) were binned to a ten-by-ten degree grid, and normalized by the surface area in each grid cell. A list of archived data sources is provided in the appendix.

Network status map:



Near line 115 "…are driven by daytime net autotrophic production (GPP + CR)". Net autotrophic production is NPP (GPP-AR).

Near line 125 "NCP (i.e., GPP + CR)". Should be GPP – CR.

Section 2.1 is well documented and informative, but it has similar sign-problems in eq 5.3 (loss processes should be negative, as in eq. 5.1), and eq. 6 (again CR should be negative). This same correction applies for Figure 3, under assumptions, "GPP + CR" should be corrected to "GPP – CR".

Thank you for these comments. As described above, we changed the notation around CR to being negative, so that NCP = GPP-CR, and NPP = GPP-AR. Equation 6 becomes:

$$T(t_1,z) = T(t_0,z) + GPP(z)\int E(t) / \bar{E}dt - CR(z)(t_1-t_0)$$

Near line 160 "…derived from particle backscatter (bbp) or beam attenuation (cp, typically at 660 nm) measurements (both m-1) using regional (e.g., Loisel et al., 2011; Cetinić et al., 2012) or global (e.g., Graff et al., 2015) algorithms.". The Graff et al, 2015 algorithm is not global, it is based on samples from an Atlantic Meridional Transect (AMT-22) and a subsection of the Equatorial Pacific.

Thank you. We clarified this statement to identify that the Graff et al. (2015) algorithm is based on a latitudinally-distributed dataset obtained from the Atlantic Meridional Transect and the equatorial Pacific (L175).

Section 2.2 for NCP. As this manuscript aims to provide a complete overview of all float-based NCP estimates/methods available, it should also include those applied to the mesopelagic layer, mostly conducted in the Southern Ocean to infer respiration rates, and thereby NCP, from oxygen drawdown:

- Martz, Todd R., Johnson, Kenneth S., Riser, Stephen C., (2008), Ocean metabolism observed with oxygen sensors on profiling floats in the South Pacific, Limnology and Oceanography, 53, doi: 10.4319/lo.2008.53.5_part_2.2094.
- Hennon, T. D., Riser, S. C., and Mecking, S. (2016), Profiling float-based observations of net respiration beneath the mixed layer, Global Biogeochem. Cycles, 30, 920– 932, doi:10.1002/2016GB005380.
- Arteaga, L. A., Pahlow, M., Bushinsky, S. M., & Sarmiento, J. L. (2019). Nutrient controls on export production in the Southern Ocean. Global Biogeochemical Cycles, 33, 942– 956. https://doi.org/10.1029/2019GB006236
- Su, J., Schallenberg, C., Rohr, T., Strutton, P. G., & Phillips, H. E. (2022). New estimates of Southern Ocean annual net community production revealed by BGC-Argo floats. Geophysical Research Letters, 49, e2021GL097372. https://doi.org/10.1029/2021GL097372

We appreciate this feedback. However, mesopelagic respiration is beyond the intended scope of our manuscript, which is meant to focus on float-based PP methods relating to autotrophic production in the upper / euphotic ocean. As we describe in the manuscript's introduction, our intention is to describe the float-based PP methods that are seen as emerging alternatives to traditional PP approaches which rely on ships or satellites. In this context, a description of processes occurring below the euphotic zone is out of scope for our manuscript.

We did, however, include some additional text in the introduction to identify that we are reviewing GPP/NCP methods for this specific purpose and noting that other recent literature (the references you provide) has presented float-based methods to evaluate NCP/respiration in the deeper water column. A modified final introductory paragraph is as follows, with revised text underlined.

The primary objective of this paper is to demonstrate the potential of autonomous platforms, exemplified by BGC-Argo floats, for expanding the spatial and temporal coverage of PP estimates in the upper ocean. This paper explores float-based approaches for estimating GPP and NCP, since those methods are more mature than emerging approaches for NPP quantification (Arteaga et al., 2022; Yang, 2021; Estapa et al., 2019; Long et al., 2021). While recent literature has presented float-based methods for quantifying PP metrics in the interior ocean (e.g., Martz et al., 2008; Hennon et al., 2016; Arteaga et al., 2019; Su et al., 2022), the focus of this manuscript is on methods that resolve processes occurring principally within the euphotic zone. To facilitate a full exploitation of these new opportunities, we take stock of the float-based tools currently available to researchers and identify their strengths and limitations. After providing an overview of the emerging float- and glider-based PP approaches, we present quantitative analyses to demonstrate the current application of these methods, as single or combined tools.

References for the NPP literature cited in the previous paragraph:

Arteaga, L. A., Behrenfeld, M. J., Boss, E., and Westberry, T. K.: Vertical Structure in Phytoplankton Growth and Productivity Inferred From Biogeochemical-Argo Floats and the Carbon-Based Productivity Model, Global Biogeochemical Cycles, 36, e2022GB007389, https://doi.org/10.1029/2022GB007389, 2022.

Estapa, M. L., Feen, M. L., and Breves, E.: Direct Observations of Biological Carbon Export From Profiling Floats in the Subtropical North Atlantic, Global Biogeochemical Cycles, 33, 282–300, https://doi.org/10.1029/2018GB006098, 2019.

Long, J. S., Fassbender, A. J., and Estapa, M. L.: Depth-Resolved Net Primary Production in the Northeast Pacific Ocean: A Comparison of Satellite and Profiling Float Estimates in the Context of Two Marine Heatwaves, Geophysical Research Letters, 48, 1–11, https://doi.org/10.1029/2021GL093462, 2021.

Yang, B.: Seasonal Relationship Between Net Primary and Net Community Production in the Subtropical Gyres: Insights From Satellite and Argo Profiling Float Measurements, Geophysical Research Letters, 48, 1–8, https://doi.org/10.1029/2021GL093837, 2021.

In response to this comment, and a subsequent one, we also included references to the Arteaga et al. (2019) and Su et al. (2022) papers in the description of published float-based NCP studies. Please see below.

Section 3 suggests that examples of GPP and NCP will be shown at local and global scales. However, a local/regional example is shown for only NCP, and a global example is shown for only GPP. These GPP and NCP analyses seem therefore disconnected between them and from the previous sections of the manuscript.

Thank you for this evaluation. Around L427 and L508 we attempted to describe why analogous analyses are not yet feasible for both PP metrics. For example, on L427-430, we state that

“*To demonstrate the current capacity for float-based PP studies at local scales, we performed a case study analysis of float/glider NCP data from OSP. A similar analysis is not presently feasible for GPP, owing to the small number of localized studies using floats and gliders, and the currently insufficient number of profiles available to conduct GPP calculations from composite diurnal cycles*”.

Presently, there have not been enough float/glider PP studies in a single region to compile those data and perform an analysis similar to our local NCP analysis. We were also unable to perform our own GPP calculations at very fine spatial scales due to the high number of profiles required to make those calculations. We clarified these point at the end of L430 by adding the following text:

Indeed, there have not been enough published float-based GPP studies to date in a single region to compile those data and perform an analysis similar to our local NCP analysis. Moreover, we could not perform our own local GPP calculations due to the high number of profiles required to make those calculations. These factors currently preclude an analogous analysis of GPP methods at localized scales.

In addition, a coarser NCP case study was not feasible because previous work evaluating NCP at basin scales is quite sparse and disparate - only a few studies (Johnson et al., 2017; Yang et al., 2019; Emerson and Yang, 2022, and the references you included on mesopelagic respiration/ANCP) have used compiled float data to evaluate NCP on coarser scales. Unfortunately, the limited number of studies preclude inter-comparisons of their results. Also, performing new global or basin scale NCP calculations in this manuscript is beyond our scope. To address this apparent inconsistency in our analyses, we added the following text to the opening paragraph of section 3.2.1, L508 (new text underlined):

Building on recent work by Johnson and Bif (2021) and Stoer and Fennel (2022), we performed new global GOP and GCP calculations using the available BGC-Argo array. We summarize those calculations here and provide further details in the appendix. Presently, a similar analysis is not feasible for NCP, as global scale NCP calculations have not yet been attempted by the community, and only a small handful of studies have calculated NCP at basin scales (see section 3.1). As a result, intercomparisons of published results at these scales are not feasible, and new calculations of global NCP are beyond the scope of the present paper.
For our GPP calculations, we followed Stoer and Fennel (2022) by compiling all available high-quality BGC-Argo ΔO2 and bbp-POC data collected between January 2010 and December 2022,

Near line 400 “ Float-based NCP studies are somewhat more numerous than GPP studies (Table A2) but are similarly limited in their geographic extent. NCP has been well-studied around Ocean Station Papa (OSP; 50oN, 145oW) in the subarctic NE Pacific (sect. 3.1.1), and only a handful of localized studies have occurred elsewhere, such as in the S. China Sea

(Huang et al., 2018) and the NW Atlantic (Alkire et al., 2014; Yang et al., 2021) (Fig. 2c)". This is incorrect, as it omits the Southern Ocean studies mentioned above.

Thank you for this feedback. We included some of Southern Ocean citations that you provided by including the following text immediately after the quoted section (L421):

"Several float-based studies have quantified ACNP in the Southern Ocean, however, that work has principally focused on processes occurring below the euphotic zone (e.g., Martz et al., 2008; Hennon et al., 2016; Arteaga et al., 2019; Su et al., 2022)"

Near line 475: " No studies to date have estimated global NCP from floats. Johnson et al. (2017) (Southern Ocean), Yang et al. (2019), and 475 Emerson and Yang (2022) (both Subtropical Ocean) have, however, provided extensive assessments of (A)NCP from a compilation of multiple floats. Johnson et al. (2017) used BGC-Argo data to characterize ANCP in the Southern Ocean by compiling NO3- data from 24 floats deployed between 2009 and 2016. Similarly, Yang et al. (2019) and Emerson et al. (2022) compiled O2 data from multiple floats to estimate ANCP in the North and South Hemisphere Subtropical Ocean.". Again, the studies listed above also used a compilation of floats to infer NCP in large regions of the Southern Ocean and should be referenced here.

Thank you. We included these references by modifying the referenced text, beginning on L496, as follows (revised text underlined):

No studies to date have estimated global NCP from floats. Johnson et al. (2017) (Southern Ocean), Yang et al. (2019), and Emerson and Yang (2022) (both Subtropical Ocean) have, however, provided extensive assessments of (A)NCP from a compilation of multiple floats. Johnson et al. (2017) used BGC-Argo data to characterize ANCP in the Southern Ocean by compiling NO_3^- data from 24 floats deployed between 2009 and 2016. Similarly, Yang et al. (2019) and Emerson et al. (2022) compiled O_2 data from multiple floats to estimate ANCP in the North and South Hemisphere Subtropical Ocean. Lastly, some recent work (e.g., Martz et al., 2008; Hennon et al., 2016; Arteaga et al., 2019; Su et al., 2022) compiled data from subsets of the Southern Ocean BGC-Argo array to quantify ANCP and respiration below the euphotic zone. Those studies, however, are out of scope for the present manuscript in which we focus on reviewing methods resolving PP metrics primarily within the euphotic zone.

Section 3.1.1. I think this analysis would be better presented in a manuscript dedicated exclusively to NCP or productivity fluxes at OSP. This way, the methodology could be better explained and expanded in a section of its own.

Thank you. Please see our response to your suggestion to divide the manuscript into separate GPP and NCP papers above. We also feel that the details on the NCP case study provided in Appendix A are sufficient for explaining the methodology employed in that analysis.

Section 3.2. Again, this section hints at the presentation of global NCP and GPP estimates, but results are presented only for GPP. This type of inconsistency could be addressed by having separate manuscripts on GPP and NCP.

Thank you. Again, please see our response to this suggestion above. Please also note our comments above describing why global NCP analyses were not performed in this manuscript.

Near line 490 “Our calculations, we extend the work of ..”. This sentence needs correction.

Thanks for catching this error. We changed the sentence from “*Our calculations, we extend the work of Johnson and Bif (2021) and Stoer and Fennel (2022)*” to “*Our calculations extend the work of Johnson and Bif (2021) and Stoer and Fennel (2022)*”

Near line 505 “ There is generally good agreement between float O2- and bbp-based GPP and between the float estimates and independent GOP estimates derived from bottle sampling (Fig. 6b,c)”. Also line 555 “Float-based GPP estimates have been shown to compare well with independent data, and well between O2- and POC-based estimates (see our global GPP case study, sect. 3.2, also Johnson and Bif, 2021; Stoer and Fennel, 2022).” I do not agree with these statements. On the contrary, I see a considerably disagreement between the zonally-averaged estimates presented in Figure 6b. From here on, most of the subsequent analyses are based on the premise of an agreement between independent GPP estimates, which is not supported by the presented analysis. The design and focus of the GPP and NCP analyses presented in section 3 do not seem to converge well together within one single manuscript. Therefore, I would strongly recommend having two different works for each topic. Overall, the review and novel analyses conducted with respect to GPP seem to be more mature than those for NCP. Perhaps the authors could consider approaching the topic of float-based GPP estimates first in a more concise manner.

Thank you for this feedback. Please note our response to your suggestion to divide the manuscript above.

We agree with your concerns about the comparison between O2 and bbp-based GPP estimates; there are important differences between these different metrics, particularly in the zonally-averaged estimates. To better address the apparent inconsistencies between O2 and bbp GPP estimates, we revised the text starting on L583 as follows (revised text underlined):

Float-based GPP estimates have been shown to compare well with independent data, and O_2 and b_{bp}-based estimates generally correlate with one another (p-value < 0.05 and R^2 = 0.47 through paired data in upper 60 m; Fig. 7). With some exceptions (e.g., surface waters between 0-30ºN) offsets between O_2 and b_{bp}-based estimates are often within the standard error of the diurnal cycle approach (Fig. 6, and see results from Johnson and Bif, 2021; Stoer and Fennel, 2022). However, when compared directly, the ratio between ΔO_2-GOP and POC-GCP is not always consistent with the expected relationships based on documented PQ and PER variability (Fig. 7). For example, given an estimated range of ~18-47% DOC production during photosynthesis (median PER value of 32.5% ± 14.4% standard deviation calculated from Moran et al., 2022), and a PQ range of 1-1.45 (Laws, 1991), the ratio between ΔO_2-GOP and POC-GCP uncorrected for PER should be between ~1.2 and 2.6 (shaded region in Fig. 7). Considering an even broader PER range of ~2-50% (global *confidence interval from Baines and Pace, 1991) results in an expected GOP:GCP ratio of ~1-2.9. Yet, in our depth-resolved, global GPP dataset, we derived a median ratio of ~3.1 ± 0.2 (median ± confidence interval) for estimates derived in the upper 60 m. When considering all depths (up to 200 m), the median ratio is ~4.1 ± 0.6, reflecting the lower signal-to-noise ratio of diurnal O_2 or b_{bp} variability at depth. For comparison, Briggs et al. (2018) calculated a ratio of ~2.6 between mixed layer O_2-GOP and c_p-GCP during a NW Atlantic spring bloom. These results imply higher PQ values and/or DOC production rates and may indicate that these terms are non-uniform across the global ocean. Using static PQ or PER values in GPP calculations (as in Stoer and Fennel, 2022 and in our global GPP case study) likely contributes to the uncertainty in the resulting GPP datasets, and partially explains the offsets we observed O_2- and POC-based GPP estimates, and differences between the float- and bottle sample GPP values. Other sources of uncertainty <u>and causes for potential and apparent offsets between O2- and POC-based estimates</u> are discussed in the following paragraphs.*

Please note that the remainder of section 4.1 describes a number of reasons for potential and apparent offsets between O2 and bbp based GPP estimates, including uncertainty in the bbp-to-POC relationship, diurnal variations in O2 or bbp not related to photosynthesis, and differences in the number and locations O2 and bpp profiles.