The authors evaluated data from two fast-profiling BGC Argo floats in the western North Pacific, investigating phytoplankton production and remineralization in response to mixing events. While the overall approach is interesting and the data source is appropriate for the research question, there are some methodological issues that will need to be addressed.

Strengths:

- Daily profiling BGC Argo floats are a good fit for addressing the research questions posed.
- The Introduction is overall well written and to the point.
- I commend the authors on a thorough discussion of whether or not events observed by the floats (i.e., the individual mixing events) can be treated as a timeseries, i.e., whether the float followed a consistent water mass or not.
- Calibration of backscatter sensor with POC before and after float deployment.

Weaknesses:
The most notable weakness, in my opinion, is the mixing of the nitrogen analysis (i.e., the decrease in nitrate over time) with an increase in POC as measured from optical backscatter. I would posit that these two can’t be directly related (i.e. to infer Redfield ratios, as was done in this study). Let’s take the mixed layer example: assuming that the Lagrangian assumption holds, then the decrease in nitrate over a short (~1 week) period of time after a deep mixing event would be related to all organic matter produced during that time, regardless of the fate of that organic matter. But the POC measured on any given day is not just the result of production, it is also affected by losses (due to grazing and export). These losses can’t be directly measured (at least not in the mixed layer), but the “lost” POC would also have contained nitrogen that is part of the “nitrogen budget” estimate. Therefore, the two measurements shouldn’t be directly compared or related, at least not in the sense that it is done in this manuscript, i.e., to infer a Redfield ratio.

There is mention of chlorophyll samples being taken on the voyage, but then the fluorescence sensor does not appear to have been calibrated. Ideally, this should be done, or there should be mention of why it wasn’t/couldn’t be done. C:Chl ratios should also not be discussed if the fluorescence sensor wasn’t calibrated.

While the beginning of the manuscript is overall well written and easy to follow, this cannot be said for section 4.2 from line 313 onwards. This section needs serious editing and revising, both in language/clarity and possibly in the approach. I can’t say I can fully follow the discussion from line 325 onwards, but it appears to me that an attempt is made to account for POC added via fragmentation of particles settling from above (?). But what about POC loss due to particles settling out of the layer of interest? Not all disappearing POC is lost due to remineralization, and only the remineralized part will be reflected in O2 and N. It is telling that the O2/N ratio observed was close to Redfield, but the C/N and C/O2 ratios were not. The first (O2/N) will reflect any organic matter that came through the depth layer under investigation and was remineralized (i.e., matter sinking down from above and being remineralized but not necessarily showing up as POC in the backscatter sensor, for example because it is too big or too small, or indeed dissolved). The POC observed in that layer (which is what presumably goes into the C/N and C/O2 ratios calculated) is likely just a fraction of all the carbon that moves through this layer, and its rate of change in a given depth layer is a combination of additions from above, sinking below, and remineralization, with only the latter reflected in the rate of change of O2 and N. Maybe this is what the authors try to allude to in the section starting at line 325, but I couldn’t quite follow the calculation relying on the Honda et al. paper. So in the very least this will need to be clarified, and quite possibly the calculation and interpretation will need to be changed. Just as was the case for the mixed layer, I don’t think the O2 and N measurements should be directly mixed with the POC measurements, so the C/N and C/O2 ratios are misleading in my opinion. And even if an attempt is made to account for the processes that cannot be measured here (such as fragmentation of particles), I doubt that this can be done with the kind of confidence required to come to the conclusion that DOC is such a strong player in the twilight zone (I can believe that in principle, but I don’t think the evidence in the data is strong enough to support this conclusion). For example, if I understand correctly and sediment trap data are being used to account for fragmentation, I would argue that the time scales of interest (likely months for the sediment trap data) and the event under investigation (which is presumably a high-sedimentation event) don’t match up, nor do the assumed intensities in particle flux.

Recommendations:
Revise the approach to Redfield ratio calculations (see weaknesses above) for both mixed layer and twilight zone.

For instances where the Lagrangian assumption doesn’t hold (i.e., the float crossed different water masses), any analyses that rely on the Lagrangian assumption should not be conducted. The discussion of water masses should come before any further inferences are made (move section starting at line 260 upwards and make it its own section in the Discussion, with a subheader).

Consider using the Redfield ratio in a different way for an additional/new focus of the paper: Based on the disappearance of nitrogen in the mixed layer, one could use the Redfield ratio to calculate the carbon that was likely produced over the given time period in question. The difference between POC measured and POC produced (based on N) could be assumed to have been exported. This would be an interesting number to discuss? See Moreau et al. (2020) for an example of the calculations/approach (their Figure 3).

Using calibrated phytoplankton fluorescence (i.e., working in units of Chl). With this, you could also evaluate the C:Chl ratio over time (e.g., after each mixing event), for example to explore light limitation and acclimation.

Revise overall structure of discussion and re-focus the manuscript as appropriate (see comments under weaknesses and also Recommendations above)

Conclusions need editing for clarity, and the section starting line 360 should be moved to the discussion. It is not good practice to introduce a new figure in the conclusions.

Check the whole manuscript for grammar/English, with particular attention to the section between lines 313 and 344.

Specifics:

- Lines 32/33: the conclusion regarding DOC in the twilight zone may need to be revised (see above)
- Line 64: “which cannot reach the mesopelagic zone” – unclear what that refers to? Do you mean “which cannot reach the mesopelagic zone by sinking”? Please clarify.
- Lines 80-83: This section doesn’t belong into the Methods. Add to introduction?
- Line 97: what time of day did the floats surface? This will be important for the use of the fluorescence data, i.e., is an NPQ (non-photochemical quenching) correction required or not?
- Lines 103-104: “…determined for each 0.01 kg m-3 of potential density…” – not sure what is meant by this? The offset for the oxygen correction is usually assumed to be constant along a profile and only drifting through time (if at all). So this is either unusual/unnecessary or poorly explained.
- Lines 129/130: The easy explanation for why there were no spikes in POC would be under-sampling? I.e., much less resolution in POC measurements relative to bbp.
- Line 131 and onwards: Did you smooth the bbp signal at all, for any of the subsequent analyses?
- Line 135: replace “seem to reflect” with “are in accordance with” for clarity.
- Line 154: “they slowed [by] for about a week”
- Line 159: insert reference for mixed layer definition
- Line 162 onwards: there seem to be 2 different mixed layers being discussed here ("the deep mixed layer" and the surface mixed layer introduced earlier) – they need to be properly referenced and distinguished; at the moment, this paragraph is confusing.
- Lines 175-190 contain quite a bit of discussion, not just results. Either revise or call the whole section "Results and Discussion" instead of "Results".
- Lines 194-202: Make a bit clearer which statement refers to which float.
- Line 198: "which was located to the east" – it’s worth mentioning that this float also saw deeper mixing for longer.
- Lines 205-214: Was an NPQ correction done/necessary? Also consider availability of nitrate (not just light) when discussing patterns of Chl increase.
- Lines 218/219: C:Chl ratios require some sort of Chl calibration for the fluorometer. Please describe what was done for Chl. Also: rather than discussing C:Chl in bulk, you could show a timeseries as this would be interesting to explore light acclimation and limitation in the phytoplankton in response to mixing/restratification.
- Line 229: Primary productivity wasn’t directly measured here, so this is a poor choice for a header; maybe replace with “derived parameters” or something like that? More structure/subheaders in the following section would also be helpful.
- Line 236: “assuming minimal lateral advection” – but what about grazing and sinking? See major comment above under “weaknesses”.
- Lines 240-245: define what you called a “storm”, i.e., which low pressure systems qualified as a storm, and which didn’t?
- Lines 288-291: not sure what you mean here, especially in respect to the observations. This needs editing/clarification.
- Lines 314-316: Definition of depth layer and time interval in question needs to be made clearer. How was that depth interval chosen? How does it relate to the mixed layer? How do conclusions change if a different depth layer is investigated?
- Line 324: Not sure what is meant by "This is likely due to the lower expected rate of decline in POC.”?
- Line 325 onwards: What is the assumption about how fragmentation and backscatter interact?
- Lines 333-335: ...“there would still be 0.02 mmol C m-3 d-1 …” – not sure what is meant by this. Should this be left behind but wasn’t measured,...?
- Line 350: replace “correcting” with “calibrating”
- Line 351/352: the word “disturbance” is rather unspecific here. Do you mean mixing event?
- Lines 360-371: This doesn’t belong into the conclusions, I suggest adding it to the discussion.
- Figure 3: It would be helpful to indicate eddy periods in the time series.
- Figure 4: The euphotic zone (according to the legend) seems to be assumed to be static at 70 m depth. This is unlikely to be true given the range in Chl concentrations, and in any case, the method for calculating the euphotic zone should be described in the methods section.
- Figures 6 and 7: Figure captions need editing.
- Tables 1 and 2: Titles should be more descriptive, e.g. indicate that the reported numbers refer to storm events, and mention what depth layer was investigated (Table 1). It should also be “Rate of change in POC/NO3-/DO” in the tables to be more specific.

References:
Moreau, S., Boyd, P. & Strutton, P. (2020). Remote assessment of the fate of phytoplankton in the Southern Ocean sea-ice zone. Nature Communications 11:3108.