**Reviewer 1:**

## General comments

This paper assesses projected changes in phytoplankton and zooplankton in three regions of the Southern Ocean as simulated by an ensemble of Earth system models (ESMs) under SSP5-8.5. The study finds three different patterns in the three regions, with trophic amplification in two of them. A new contribution of this study is examining if there is seasonal trophic amplification, instead of just a difference between a future time period (e.g. 10-30 yr mean) and the contemporary time period. The study also attempts to attribute the changes using a budget analysis for both phytoplankton and zooplankton. Though the results are supported by evidence, not all the interpretations of the results are substantiated. The key message is important and clearly described, highlighting seasonal changes and how they could be obscured by only examining annual means. It also prompts next steps for assessing downstream impacts of the seasonal trophic amplification.

Thank you for spending time in reviewing the paper and for your thorough and constructive evaluation. Your comments have helped us strengthen the analysis, sharpen the presentation, and improve the overall clarity of the paper. Below we address each comment individually. Reviewer comments are in black, our responses in blue, and planned changes to the manuscript in *italics*.

The Results and Discussion often hypothesize how temperature, light, and nutrient availability caused the changes in plankton biomass, concentrations, and rates (e.g. L197, L280, L326). However, these terms are readily available from the ESMs on the ESGF server. A seasonal analysis of temperature, sea ice, downwelling radiation, light limitation, N limitation, and Fe limitation akin to the MLD analysis in Figures 4-6 would greatly strengthen the paper by attributing the mechanisms rather than hypothesizing on them.
The Arteaga et al. 2020 reference is not adequate for these statements because it examined Argo observational data, not model output.

Section 3.2.3 head reads “Spring and summer light availability drives positive trophic amplification in the SIZ,” but light availability is not directly examined.

We appreciate this suggestion to strengthen the link between environmental variables and plankton biomass changes. We investigated the CMIP6 limitation diagnostics and found that only three models in our ensemble (CESM2-WACCM, IPSL-CM6A-HR and UKESM1) provide the full set of light, nitrogen and iron limitation terms under all three experiments used in our study (piControl, historical, SSP5-8.5). In addition to limited model availability, the light limitation diagnostic (limirr) is confounded by the temperature effect on growth in some models. For example, in MARBL (used by CESM2-WACCM). We therefore cannot isolate the light effect from the limirr diagnostic across models. Instead, we add an analysis of the underlying environmental variables driving limitation/impacting growth, namely: sea surface temperature (tos), mixed layer depth (mlotst), surface shortwave radiation (rsntds), surface nitrate (no3os), dissolved iron (dfe) and sea ice concentration (siconc). These variables are available for almost all models and provide diagnostic potential with respect to the changes of individual physical and bgc drivers and whether they are consistent with our interpretation. We will also go through the references and lines of text as suggested to make sure they are accurately reflect the analysis presented.

### Specific comments

Additional references

L24: Add also Petrik et al. 2020, Guibourd de Luzinais et al. 2023.

L34: Add also Asch 2015, Chen et al. 2025.

L56: Suggest adding Kearney et al. 2021 that has a detailed description of the BCG model components.

We appreciated the additional relevant references. They will be added as suggested.

Methods need more detail

L53: Please list all the variables you used for this analysis and their CMOR names too.

We will clarify that the selection of the models is based on the availability of several main variables: mixed layer depth (mlotst), phytoplankton (phyc) and zooplankton (zooc) biomass concentrations. Further analyses also refer to primary production (intpp), zooplankton grazing (graz), sea ice extent (siextents), sea surface temperature (tos) and now also, as suggested, the underlying environmental variables driving limitation/impacting growth.

Section 2.1: Add definition of years for "contemporary" and "future" analyses.

Thanks for pointing out the missing definition. We have added the definitions to the first sentence in the section:

*"We use the output of a Multi-model ensemble of 12 Earth System Models (ESMs) to diagnose contemporary (2000-2009) and future (2090-2099) plankton changes in the Southern Ocean."*

Table 2: Which years?

We will add the time period of the products and model simulations in the table.

L144-147: I do not fully understand these methods. i) I don't understand why the normalization step is needed. ii) Is there only one budget per term (P, Z) from the multimodel mean, not one budget per term per model? iii) Seems like the averaging could lead to a mismatch of loss/gain terms, which then would be absorbed by the Other term.

We will expand the methods description to clarify each point:

1) The 12 ESMs in our ensemble differ substantially in their absolute plankton biomass and rate values — in some cases by an order of magnitude (cf. Figure A2). If we simply averaged the raw seasonal cycles across models, the multi-model mean would be dominated by the few models with the largest absolute values, effectively obscuring the seasonal patterns of the remaining models. To give each model equal weight in determining the *shape* of the seasonal cycle, we first normalise each model's monthly values relative to its annual mean. These normalised relative seasonal cycles for all models can be meaningfully averaged across the ensemble. We then recover absolute values by multiplying this mean relative seasonality by the multi-model mean annual value.

2) The budget terms (PP, GRAZ, O_phy, O_zoo) are computed for each individual model first. Each model's seasonal cycle of each term is then normalised by its annual mean, as described above. The normalised seasonal cycles are averaged across models, and then scaled back to absolute units using the multi-model mean annual value of each respective term. Thus, every model contributes equally to the seasonal *shape* of each budget term, while the absolute magnitudes reflect the ensemble mean. We will clarify this in the revised Section 2.

3) This concern would apply if the Other term were normalised and rescaled independently as the other budget terms. However, that is not our procedure. Partly to avoid introducing such a mismatch, we adopt the normalisation approach described in (i): each budget term is first normalised by its annual mean, averaged across models, and then multiplied back by the multimodel mean annual value of that same term. Because every term is

scaled back to absolute units before the budget is assembled, and the Other term is then computed as the residual needed to close the budget, closure is exact and the normalisation procedure should not introduce any mismatch.

Not clear where many of the numbers in the Results section come from

L206-207: How is the 67% more calculated? 10%/6%?

Yes, exactly. 67% is calculated as (zooplankton-phytoplankton)/phytoplankton: (10%-6%)/6%

L233: 30% looks closer to 20%.

Corrected.

L252: Is 50% stronger the comparison between 16% vs. 10%?

The “50% stronger” compares the trophic amplification ratio during the amplified period (Nov-Jun) to the annual mean. During Nov-Jun, zooplankton declines 14%, while phytoplankton declines 5%, giving an amplification ratio of 2.8. On an annual mean basis, the subset of models with flux output (used for the seasonal analysis in Figure 4) shows zooplankton declining 11% while phytoplankton declines 6% (thick horizontal lines in Figure 4c, d), giving a ratio of 1.83, consistent with the full ensemble values (10% vs. 6%, thin horizontal lines in Figure 4c, d). The seasonal signal is therefore (2.8 − 1.83)/1.83 ≈ 53%, rounded to “approximately 50%". We will clarify the text to make the comparison more transparent.

*“During the amplified period (red shading in Fig. 4), zooplankton biomass decreases nearly three times as much as phytoplankton (14% vs. 5%), representing a seasonal trophic amplification that is approximately 50% stronger than the annual mean trophic amplification (11% vs. 6% for the flux-output models, thick horizontal lines in Figure 4 c,d).”*

L313-314: How is the >4x quantified? The relative change in Fig 6 c and d looks like 70% vs. 35% (~2x). Is it referring only to the Jan diff that is ~4x?

Thank you for catching this. After rechecking, the amplification ratio during October-January (zoo +35.8% vs. phyto +14.6%, ratio of 2.5) is similar to the annual mean ratio (20.1% vs. 8.6%, ratio of 2.3, think horizontal lines in figure 6 c, d). The “>4 times” statement was incorrect. We will revise the text accordingly.

*“During the amplified period from October to January, both phytoplankton and zooplankton biomass increases are largest (+15% and +36%, respectively), with zooplankton increasing approximately 2.5 times as much as phytoplankton, similar to the annual mean trophic amplification.”*

L317: “eleven times as much” Can you please give the two numbers for phytoplankton and zooplankton for verification of this quantity?

We will add these values to the revised text for transparency.

*“The core of the amplified period occurs in January, during austral summer, when zooplankton biomass increases by +34% compared with +3% for phytoplankton, approximately eleven times as much and roughly five times the annual mean trophic amplification.”*

Imprecise language

L112-113: “Not capture fully” is imprecise wording. These could be captured, but ESM modelers choose not to. Omnivory occurs when >1 zooplankton type, which is represented in your ensemble. Slower turnover rates can be accomplished with differential temperature-dependence, mortality, and/or representation of life history stages.

We will revise the sentence to clarify that these are simplifications in current model design rather than inherent technical limitations.

*“These distinctive ‘top-heavy’ features reflect complex ecological and biogeochemical dynamics that are simplified in current climate model representations of lower trophic levels.”*

L270: Is a deltaT of <0.3 mmolC/m2/d sharp? What is the context or criteria for calling this sharp?

We have removed the word "sharp" as it was only meant as descriptive but not quantitative.

L285: The MLD difference looks biggest to me (from the two lines) in Jul-Sep, but the bars are the smallest then.

Yes, the absolute MLD difference between the future and contemporary periods is largest during austral winter. However, the MLD itself is also deepest during those months. The bars in the figure show the MLD change relative to the contemporary (2000–2009) mean for each respective month, not the absolute difference. Because the winter MLD is so deep, a large absolute shoaling translates to a small relative change, whereas a moderate absolute shoaling during months with a shallower contemporary MLD produces a larger relative change. We notice that the caption was incorrect, and we will correct this in the figure caption:

“Absolute seasonality (left axis, present: solid line, future: dashed line) and relative change with respect to the respective month's contemporary (2000–2009) mean (right axis, bars) of (a) mixed layer depth (MLD) …”

L292: The statement “increasing loss terms offset the higher grazing intake across all seasons” is not accurate. Losses < Grazing Jun-Oct.

Our intended meaning was that the increase in loss terms largely offsets the increase in grazing intake across the year, resulting in only a small net change in zooplankton biomass. However, the original wording overstated this, as grazing gains do exceed losses from June to October. We have revised the text accordingly.

“Although zooplankton benefit from improved grazing conditions year-round, the increase in loss terms largely offsets the higher grazing intake, resulting in only a small net change in zooplankton biomass seasonality (Fig. 5d,f).”

L295: You say there is no amplification window, but It seems that S, O, and D meet the criteria.

Our identification of amplification windows requires a trophic amplification signal in the annual mean trend. In the SPZ, neither phytoplankton nor zooplankton show a substantial annual mean change, and the small changes are not consistently in the same direction, so there is no annual mean trophic amplification signal from which to identify a seasonal window. We will also add the prerequisite to the amplification period definition:

“In regions where trophic amplification is present in the annual mean signal, we define the amplified period as the months in which the monthly relative change of zooplankton biomass over the century exceeds both the annual-mean relative change and the corresponding monthly relative change of phytoplankton.”

L322: “throughout most of the year” Just Aug-Jan, that's half not most.

We have now corrected it to half of the year.

L329: Elevated grazing only compensates Aug-Jan.

The dot-dash line in Fig. 6f shows the change in monthly tendencies, not the absolute zooplankton biomass change over the century. The elevated grazing intake leads to an overall zooplankton biomass increase from September to May (Fig. 6d), not only August to January. We have clarified this in the text.

*"The elevated grazing intake more than compensates for the higher non-attributable losses (Fig. 6f), leading to an overall increase in zooplankton biomass from September to May (Fig. 6d)."*

L330: “increase in zooplankton biomass” Losses seem to exceed grazing in Fig 6f (negative deltaT), but the change in biomass is positive Feb-May.

We appreciate this observation. Figure 6f shows ΔTzoo = $dzoo_{2090s}/dt - dzoo_{2000s}/dt$, which is the difference in the tendencies of biomass change between future (2090s) and contemporary (2000s). A negative ΔTzoo in February–May indicates that zooplankton biomass either increases slower or decreases faster in the future compared to the contemporary period during those months, but it does not imply that future biomass is lower than contemporary biomass. The same applies to the strongly positive ΔTzoo from September to January, which means that zooplankton biomass increases faster or decreases slower in the future during those months, and this accumulated surplus carries into February–May, sustaining a positive biomass change (Fig. 6d) even as the rate advantage diminishes.

L401: “different … directions” All the shifts look earlier to me in Supp Fig C3.

In Figure C3a, the initiation of the phytoplankton bloom shifts later, as indicated by the blue shading. Although the shift is small, it does show a different direction of change compared to the other phenological shifts

L94-95: Could the difference between Argo and BGC-Argo be the time range? e.g. Argo has 10 years under cooler conditions and is missing 3 warmer years?

This is an interesting point. The Argo mixed layer product covers 2000–2021, while the BGC-Argo data span 2010–2024, so there is partial temporal overlap. While we cannot rule out that the different time ranges contribute to the discrepancy, we consider it more likely that the difference reflects the sparser spatial coverage of BGC-Argo, particularly in the seasonal ice zone, as noted in the text. The BGC-Argo network remains relatively young and has fewer floats in high-latitude and coastal regions compared to the standard Argo array.

L165-166: I understand this in theory, but the figures do not meet my expectations given these explanations. e.g. Fig 4d,f: deltaZ is positive Jan-Jul, suggesting the difference between 2000s and 2090s is decreasing those months, but the relative change in those months is very large.

Thank you for raising this. Your understanding is correct. Taking Fig. 4d,f as an example: ΔT_zoo is positive from January to July, indicating that the difference between the 2000s and 2090s is indeed shrinking during those months. This is visible from the solid and dashed lines in Fig. 4d, which converge towards July. The relative change (bars), however, is calculated as the absolute difference in each month divided by the present-day monthly mean. So although the absolute difference in, say, April is smaller than in January, the present-day monthly mean biomass in April is also smaller, resulting in a similar level of relative change.

L196: Could the stability simply be that phytoplankton biomass has decreased proportionally with PP and GRAZ, such that u 〓 PP/phy and g 〓 GRAZ/phy are the same?

That is indeed what happens in the STZ. PP, GRAZ and phy all decline in near-proportion, leaving μ and g largely unchanged. However, there is no physical reason why they must decline proportionally, as PP depends on light and nutrients, GRAZ depends on zooplankton biomass and prey density, and phy responds to both production and losses. The near-proportional decline is itself a physical result, not a mathematical inevitability. As discussed in the text, it arises because mixed-layer shoaling has opposing effects on light and nutrient limitation that largely cancel in μ, while higher temperatures partly compensate for the lower prey concentration in g.

L248: October is also the biggest decrease in MLD in future compared to contemporary.

We have revised the text to note that October also coincides with the largest projected MLD shoaling, strengthening the link to intensified nutrient limitation.

“*The most pronounced phytoplankton decrease occurs in October, which also coincides with the largest projected MLD shoaling, intensifying nutrient limitation as biomass is near its seasonal*

*peak, whereas zooplankton show their greatest decline in austral summer (January and February), when the MLD is seasonally shallowest (Fig. 4d).”*

L260: I think this should say “decreasing grazing mortality.” dGrazing is lower Oct-Jun as indicated by the bars. dGrazing is negative in the amplified period, both in the budget (Fig 4c) and Dec-Apr in the g/g2000s comparison (Fig 4b), so it must be the decreased growth in (Fig 4b) that resulted in reduced PP.

The text refers to grazing mortality rate g (= GRAZ/phy, Eq. 4), not the grazing flux GRAZ itself. While ΔGRAZ is indeed negative during the amplified period (Fig. 4c), this reflects the decline in phytoplankton biomass rather than a weakening of grazing pressure per unit biomass. As shown in Fig. 4b, $g/g_{2000}s > 1$ in November, May and June, indicating that increased per-capita grazing mortality contributes to the phytoplankton biomass decline during those months. From December to April, however, the biomass decline is driven by a decreasing growth rate ($\mu/\mu_{2000}s < 1$, Fig. 4b), consistent with deteriorating growth conditions. We will specify this in the revised text as “increasing grazing mortality rate (g)” and refer to figure 4b by the end of the sentence.

Section 3.2.2: Could also mention that both phytoplankton and zooplankton show phenological shift earlier.

After consideration, we believe this information is best incorporated into the Discussion where we address phenological shifts. We will revise the text:

“Our MME projects shifts in the timing of phytoplankton and zooplankton blooms across all three zones, with both phytoplankton and zooplankton shifting earlier across most of the Southern Ocean, and only phytoplankton in the STZ showing a slight delay in bloom initiation (Fig. C3). The different degrees and even directions of phenological shifts for zooplankton and phytoplankton themselves constitute a form of trophic decoupling (Yamaguchi et al., 2022), raising the prospect of widening trophic mismatches between successive levels of the food web (Edwards and Richardson, 2004).”

L285-286: How does one see that “From August to October, enhanced phytoplankton production driven by mixed-layer shoaling induced light improvement and rising temperatures leads to a subtle increase in phytoplankton biomass” in 5e? It does not show MLD or light. And deltaT is positive for May-Sep, but very negative for Oct.

We agree that the mechanistic attribution to mixed-layer shoaling-induced light improvement and temperature increase cannot be read directly from Fig. 5e, which shows only the budget terms (ΔPP, −ΔGRAZ, −ΔOphy). We will examine the limitation factors, as suggested in the general comments, to verify the underlying mechanism. Regarding the reviewer's second point, we note that ΔTphy (dot-dash line in Fig. 5e) represents the difference in the tendency of biomass change between the future (2090s) and contemporary (2000s) periods, not the absolute biomass change. A strongly negative ΔTphy in October therefore means that biomass increases much more slowly in the future compared to the contemporary period, but does not imply that future biomass is lower than contemporary biomass overall (Fig. 5c).

L304: Could also mention a phenological shift earlier for both phytoplankton and zooplankton.

As mentioned above, We believe this broader context fits more naturally in the Discussion.

L327: Grazing can also be enhanced by warming. This goes for all regions.

We agree. Warming can enhance grazing rates in some models, and this effect is implicitly captured in the GRAZ term of the zooplankton budget. We will revise the text to acknowledge that temperature increases may also contribute to enhanced grazing.

*"The increase in zooplankton biomass can be attributed to enhanced grazing (Fig. 6f), driven by greater food availability from higher integrated phytoplankton biomass (Fig. C1), by higher*

*predator-prey encounter rates resulting from elevated phytoplankton concentration, and by warming-induced increases in grazing rates."*

Supp Fig B3: Point to your Supp figure with surface concentrations anytime you mention enhanced concentrations rather than integrated biomass in Results.

We will add the reference to the supp figure as suggested.

Results: How is surface irradiance projected to change in these 3 regions? Other studies (e.g. Krumhardt et al. 2022) have shown that some get cloudier and others less cloudy?

We will add analysis of surface shortwave radiation (rsntds) in the results to show exactly how surface irradiance is projected to change, alongside referencing existing studies.

L350-353: You say your results emphasize the ZPC mechanism as if the other two were not important components of your results, but you didn't assess ZGE or ZTL. Also, ZPC is GRAZ/PP right? Might be good to show in your 3 regions historically and over time. Especially since the ZGE, ZTL, and ZPC changes were quite variable around the S Ocean in the GFDL model.

We agree that our study primarily assesses ZPC (the ratio of grazing to primary production) rather than ZGE or ZTL, and we have revised the text to acknowledge this more clearly. Diagnosing ZGE would require zooplankton secondary production output, which is not routinely available across the MME, and ZTL is difficult to assess given the limited number of zooplankton functional types in most models (Table 1). We will add a figure showing ZPC (grazing/PP) across the three regions over the century to support this discussion.

*"The three key food web properties identified by Stock et al. (2014), zooplankton growth efficiency (ZGE), the trophic level of zooplankton (ZTL), and zooplankton-phytoplankton coupling (ZPC), remain relevant here. While ZGE and ZTL were not explicitly assessed in this study, our findings suggest that ZPC acts as a proximate driver of seasonal trophic amplification in both the STZ and SIZ, consistent with the seasonal analysis of Xue et al. (2022a)."*

L359: You can see if this is true of the heterotrophic processes by checking the temperature-dependence terms in each BGC model. Some examples in Kearney et al. 2021.

We have looked into the temperature-dependence terms for heterotrophic processes across our model ensemble using the Q10 values reported in Kearney et al. (2021, Table 5). Among our MME, only PISCESv2 (CNRM-ESM2-1) parameterizes zooplankton grazing with a stronger temperature dependence than phytoplankton growth (Q10 = 2.14 vs 1.90). COBALTv2 and MARBL apply the same Q10 to all processes (1.88 and 1.70, respectively), while the remaining models (HAMOCC, MEDUSA, WOMBAT, CanOE, MOPS) treat zooplankton grazing as temperature-independent (Q10 = 1). This suggests that in most models of our ensemble, warming-enhanced grazing is not a dominant mechanism driving trophic amplification. We will add a brief discussion of these model differences in temperature sensitivity and cited Kearney et al. (2021).

"*Yet, most models in our ensemble treat zooplankton grazing as temperature-independent or apply the same Q10 to all biological rates, with only PISCESv2 parameterizing a stronger temperature sensitivity for grazing than for phytoplankton growth (Kearney et al., 2021).*"

L360-361: Could you explain this "partially counteracted amplification" more? Are you only considering increased grazing with increased temperature (not increased metabolic demand and other losses)? e.g. Increased phytoplankton + Disproportionately higher grazing = Greater increase in zooplankton (reinforce positive amplification), whereas Decreased phytoplankton + Higher grazing = Weaker decrease in zooplankton (counteract negative amplification). This would only be true if the increased grazing on less phytoplankton biomass can meet the higher metabolic demands.

Archibald et al. (2022) did consider both zooplankton grazing and respiration in their differential temperature sensitivity experiments using the MIT Darwin model, and found that warming increased heterotroph-to-autotroph biomass ratios even accounting for higher metabolic costs. However, as noted in our revised discussion above, most models in our ensemble treat zooplankton grazing as temperature-independent, meaning that this mechanism is largely absent from our projections. Therefore, the omission of differential temperature sensitivity is unlikely to introduce substantial bias in our conclusions.

L372: It could also arise from biomass not accumulating in the zooplankton groups because their generation time is shorter than the zooplankton with complex ontogeny that actually live in this region (and thus biomass turnover rate is faster).

This is a plausible contributing factor. The zooplankton groups represented in most CMIP6 BGC models have simplified life cycles and relatively short generation times compared to the complex ontogeny characteristic of high-latitude zooplankton. Zooplankton parameterizations in ESMs are typically calibrated towards global mean conditions, which may not adequately represent the unique characteristics of high-latitude zooplankton communities. The faster biomass turnover rates in the models may therefore prevent zooplankton biomass from accumulating to the levels observed in nature. We will add this limitation in the revised manuscript:

*"This systematic bias likely reflects an underestimation of top-down control in current Earth system models, which may arise from the limited number of zooplankton functional types represented (Table 1), insufficient representation of zooplankton omnivory (McCauley et al., 2018), or simplified zooplankton life cycles that do not capture the complex ontogeny of high-latitude species."*

L377 and Fig C4: I would like a statistical test of this if possible. I notice some patterns. i) 2p2z-noFe outlier in STZ and SPZ. ii) 1p1z low range of STZ and high range of SPZ. iii) withFE lower than noFe in SPZ.

L383-384: How uncertain are the other physical changes, like sea ice, across the MME?

Sea ice projections show considerable spread across our multimodel ensemble (Fig. B5), consistent with the large intermodel uncertainty reported for Antarctic sea ice in CMIP6 (Roach et al., 2020).

L400: The phenological shifts can also be seen in Figs 4-6.

We will include the reference to Figs 4-6.

Technical corrections

L29: The Asch et al. 2019 paper was specifically about fish spawning. This detail should be added.

We will revise the text to specify that Asch et al. (2019) refers to fish spawning phenology.

*"Different trophic levels are governed by different seasonal drivers, with phytoplankton responding to more light and nutrient availability (Doney, 2006) and fish, particularly their spawning, responding more to temperature and food availability (Asch et al., 2019)…"*

Figure 1: Add a colorbar with units for chl. In the SIZ, the arrow is at the start of the seasonal cycle, which disagrees with the text saying, "austral spring and summer," while the biomass is higher than the historic pattern in the last half of the cycle.

We will add a colorbar with units for chlorophyll to Fig. 1a. We will also adjust the arrow position in the SIZ panel to better align with the austral spring and summer period described in the text.

Table 2 and throughout: I'd suggest using mgC instead of mmolC for plankton biomasses, as I think plankton ecologists are more familiar with thinking in these units.

We appreciate this suggestion. However, we decide to keep mmol C as this is the native unit of CMIP6 model output and is standard across the Earth system modelling community, which ensures direct traceability to the source data.

Figure 2: i) Please make this figure the whole page width and choose a different color for the purple line. It is too similar to the black. ii) I also suggest using the same y-axis limits for all left-hand phytoplankton and zooplankton, 0-180, and the same y-axis for all right-hand phytoplankton and zooplankton, 0-360. iii) Clarify if both the red and purple lines use the left y-axis labeled "obs".

The current figure size follows the journal template for two-column figures. We have changed the purple line to a more distinct colour to improve visibility against the black line. (ii) We intentionally use different y-axis limits across panels so that the relative seasonality in each zone is visible despite the large differences in absolute biomass. Using a uniform axis range of 0–180 or 0–360 would compress the seasonal cycle in zones with lower biomass, obscuring the patterns we aim to highlight. (iii) We have clarified in the caption that all coloured lines (observational products) use the left y-axis.

L120: Can you add an estimate of how much zooplankton data is overlapping in both of these vs. how much is different? I assume <2012 is the same as COPEPOD and >2012 is new to Yang?

The Yang et al. (2022) dataset comprises 2,909 mesozooplankton stations spanning 1932–2020. Of these, 1,125 stations (~39%) come from the earlier Moriarty and O'Brien (2013) database, which shares common origins with the COPEPOD database (O'Brien, 2007). The remaining 1,784 stations (~61%) were newly compiled from published papers, cruise reports, and unpublished data. The overlap is therefore substantial for the earlier records, though the split is not strictly temporal. We will revise the text to indicate the difference:

"*Modelled integrated zooplankton biomass is validated against monthly integrated mesozooplankton biomass from two observational datasets: the COPEPOD database (O'Brien, 2007) and the dataset compiled by Yang et al. (2022), which includes 2,909 stations spanning from 1932 to 2020, with approximately 61% of stations newly compiled beyond COPEPOD (Fig. 2g–i).*"

Eq 6: Technically, there is an efficiency constant on the GRAZ term such that not all GRAZ becomes zooplankton biomass.

We thank the reviewer for this correction. Indeed, only a fraction of grazed phytoplankton is assimilated into zooplankton biomass. We will revise Eq 6 to:

$$Tzoo = dzoo/dt = \varepsilon \cdot GRAZ - Ozoo \quad (6)$$

where ε is the model-specific zooplankton assimilation efficiency. All values and references will be listed in the new Table in the appendix.

Figure 4b, 5b, 6b: Add the one-to-one line to these plots. Also, how is one supposed to be able to distinguish April/Aug, Jan/June/July and Mar/May?

Good point. We will add the 1:1 lines to these scatter plots and we will consider replacing the letters with numbers to indicate the months.

L273 and throughout: "predator-prey" is the common terminology, not "prey-predator."

Thank you for pointing this out. We have replaced "prey-predator" with "predator-prey" throughout the manuscript.

Figure A1: It would be nice to also see a second figure next to it that shows the # of months with data for each location.

We will add a new figure (Figure R1.1) to Figure A1 showing the monthly distribution of Argo profiles used in this study across the three zones. This allows the reader to assess the temporal sampling density and identify potential seasonal biases in observational coverage.

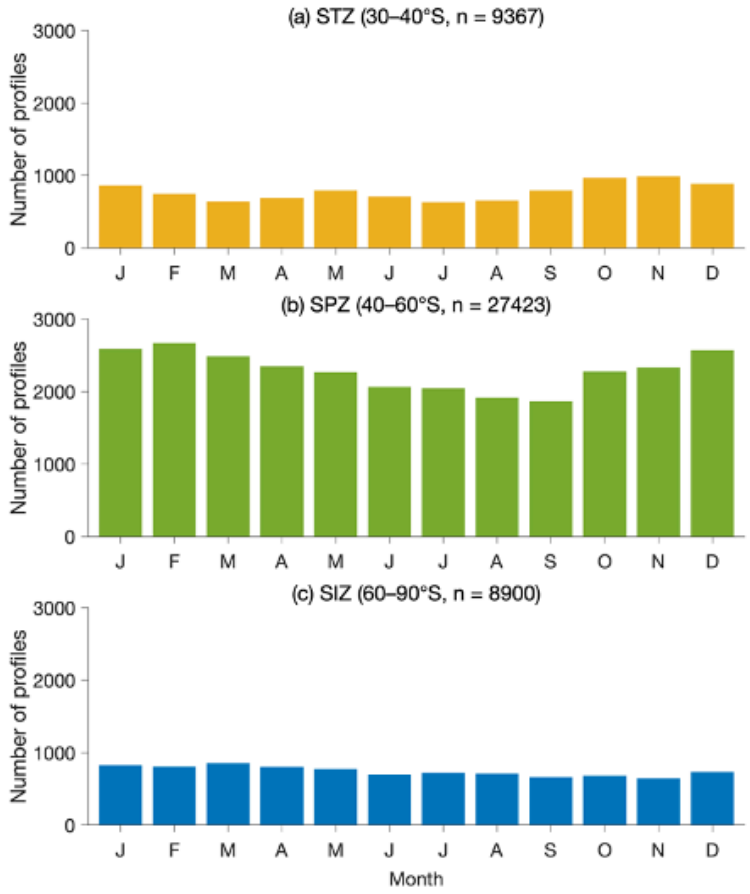



Figure R1.1: Monthly distribution of the Argo profiles used in this study (n = 45,690 profiles from 355 floats) in three latitude bands: (a) the Subtropical Zone (STZ, 30–40°S; n = 9,367), (b) the Subpolar Zone (SPZ, 40–60°S; n = 27,423) and (c) the Seasonal Ice Zone (SIZ, 60–90°S; n = 8,900). Bars give the number of profiles collected in each calendar month.

Figure A2: Label (a)-(l) with each ESM so one does not have to go back and forth between the table and this figure to know which is which.

The figure will be adapted as suggested.

Figure B3: Suggest adding water column-integrated plots next to these to make comparison easier. Or add dashed lines to these.

Fig. B3 is intended to show changes in surface concentration to support the mechanistic interpretation in the link between surface phytoplankton concentration and zooplankton grazing), rather than for comparison with integrated biomass. The integrated values are presented separately in Fig. 3.

L464/Figure B4: Just curious, why is the Fe time series so cyclical?

The multidecadal fluctuations in the surface iron time series, particularly in the STZ and SPZ, are mostly due to large swings in CESM. However, it is not clear to us and beyond the scope of this study why the CESM shows that large internal variability.

Figure B5: Would these changes be isolated to only the SIZ?

The sea ice extent shown in Fig. B5 is the total Southern Ocean sea ice extent (siextents) as output by the CMIP6 models and is not restricted to the SIZ. However, since sea ice cover is predominantly located south of 60°S, the projected changes are most relevant to the SIZ.

Figure C1: Why surface and not integrated here?

Integrated phytoplankton biomass seasonality is shown in panel c of figure 4,5,6.

Figure C3: I do ot see any blue bars. Is that correct?

In Figure C3a, the initiation of the phytoplankton bloom shifts later, as indicated by the blue shading.

**Reference:**

Roach, L. A., Dörr, J., Holmes, C. R., Massonnet, F., Blockley, E. W., Notz, D., ... & Bitz, C. M. (2020). Antarctic sea ice area in CMIP6. *Geophysical Research Letters,* 47(9), e2019GL086729. https://doi.org/10.1029/2019GL086729