Reply to the reviewers (2nd review) Saavedra-Pellitero et al.; ms re-submitted to EGUSphere

Dear Authors,
Thank you for submitting your revised paper and for responding to the reviewer comments. Both reviewers have now read your revised manuscript and both have noted the improvements that you have made to the work. However, both of them have also identified some small improvements that can be made, where you have not quite addressed some of their original points.

So I would encourage you to submit a further revision and responses to their remaining points (the details of these are available in the second review reports that both have submitted). After you have done this, and assuming you address their remaining points, I envisage accepting the paper for publication in the journal.

I look forward to receiving you final revised paper and the accompanying responses to the second set of reviewer comments.

best wishes
Jamie Shutler

General comments:

This reviewer acknowledges that the authors have made an effort to tighten-up the manuscript from the first version and scaled back their conclusions accordingly. I note that due to the tightening of this manuscript, the stated impact of this paper has also been scaled back but, at least, it is more realistic now. After re-reading the paper, there are still a few issues to attend to that I will discuss below.
Dear reviewer, thank you very much for carefully checking the new version of our manuscript and for your suggestions. Your time and effort are greatly appreciated.
Your comments definitely helped us to provide an improved version of the initial manuscript.

The title could be honed even more to state exactly what you are showing in this paper. I would suggest something like “Mismatch between coccolithophore-based estimates of particulate inorganic carbon (PIC) concentration and satellite-derived PIC concentration in the Pacific Southern Ocean.” I think this is more realistic about what this paper is actually about.
We modified the title accordingly. The suggested title is more realistic about the contents of this piece of research.

The revised paper explains that they used different types of satellite data to estimate PIC remotely, (level 2 and level 3). Each type of satellite data has a different space and time averaging. They should state emphatically that they compared these different time and space averages of satellite data to the same ship data, (always collected at the same space/time scales). Note, these different space and time averages of the two data sets will affect any mismatch between the ship and satellite results, as well as the potential errors.

Obviously, a lot can change in a coccolithophore population in +/- 8d or even +/- 1 month (see line 246) from a ship sample! It is still not really clear why they didn't simply use the level 2 data only (which would have less time and space errors but admittedly lower numbers of possible matchups due to the extreme cloudiness of the region). Did the authors ever use level 2 and level 3 satellite data in the same comparison with ship data? If so, those analyses cannot be used.
Due to frequent cloud cover in the Southern Ocean, the availability of daily MODIS L2 data was limited, making it impossible to obtain matches for each field sampling location within the typical ±3-hour window (Bailey and Werdell, 2006) or on the same day. To address this lack of suitable data , we opted to aggregate all available Level 2 data for each location over the entire duration of the field campaign ± 7 days. We believed this allowed for a more robust comparison of coccolith-estimated PIC with satellite data, given the limited number of available daily observations.

Regarding the use of L3, we included these data to supplement the MODIS daily values due to their poor availability. MODIS L3 product's broader spatial and temporal resolution helped address L2 data gaps and ensured that the comparison of lab-measured PIC and satellite-derived PIC was meaningful despite the limitations of the satellite-derived data. The fact that L3 products are derived directly from L2 data ensures consistency in the underlying scientific methodology, despite the differences introduced by the spatial and temporal averaging in L3.

As previously explained, we did not conduct a formal validation in this study, but instead focused on comparing trends across coccolith-based and satellite-derived PIC. However, it is important to note that L3 products are derived directly from L2 data, remaining consistent with the underlying scientific methodology and retrieval processes (Scott and Werdell 2019). Therefore, while the aggregation in L3 can lead to smoothing and the loss of finer-scale variability, L2 validation results are generally applicable to L3 data (Scott and Werdell 2019).

In our study, we extracted values from both L2 (daily) and L3 (8-day and monthly) data for the same field locations. For L2, we followed the method outlined by Bailey and Werdell (2006), while for L3, we used the value of the enclosing grid cell. We then compared these point values, i.e. L2 extracted point values and L3 extracted point values, to the actual field measurements (figures 3 and 4). We were fully aware that this approach did not allow for a 'proper' match-up, due to the unavailability of L2 data. Therefore, we focused on comparing trends within what we considered a reasonable temporal window, understanding that while the data available did not allow for a strict validation, it still provided valuable insights into the general patterns of variability across both data types. This approach allowed us to assess broader trends while being mindful of the limitations and uncertainties introduced by the aggregation of L3 data.

Therefore, the use of both L2 and L3 data in our study was driven by practical considerations, such as the scarcity of available L2 data due to cloud cover and the need for more robust data for comparison with field measurements. The decision to include L3 data allowed for a more comprehensive comparison of satellite-derived PIC. While this introduces additional uncertainties, these variations are accounted for in our interpretation and do not diminish the overall consistency between L2 and L3 data.

We recognize the potential limitations in comparing in situ point measurements with satellite-derived 5x5 pixel averages, as this assumes a certain degree of spatial homogeneity that is not always present. However, this approach is standard in remote sensing studies (Bailey and Werdell, 2006), and despite inherent variability in the ocean, it remains a practical solution for comparing field-based and satellite data. In our study, this approach was further refined by aggregating L2 data over the full field campaign period ± 7 days to account for temporal variations, and we carefully considered the potential discrepancies when interpreting our results.

Given the unique challenges of working in the Southern Ocean, where cloud cover is frequent and can limit satellite data availability, we proceeded with the study using the available data. While the availability limitation of daily satellite observations posed limitations, we believe that aggregating the available L2 data and supplementing the comparison with weekly and monthly satellite observations (i.e. L3 data) was a reasonable approach to attempt a more comprehensive view of PIC variability across measuring techniques. These efforts were essential in overcoming the practical challenges of finding suitable remote sensing data in this region.

We have not combined L2 and L3 data in any single analysis. Each dataset (L2 and L3) has been analyzed independently, and the results for L2 and L3 data were not directly compared to each other for any given sample location.

Reference: Joel P. Scott and P. Jeremy Werdell. Comparing level-2 and level-3 satellite ocean color retrieval validation methodologies. Opt. Express 27, 30140-30157 (2019).

The estimates of PIC per coccolith that they cite regarding their model are huge. Beginning with lines 338-356 they discuss published estimates of calcite content per coccolith for E. huxleyi. They mention 0.015pmol per coccolith (Charalampopoulou et al., 2016), then they cite the author's own estimate of 4.64 pmol per coccolith as well as the calcite per coccolith used in this paper (of 1.66pg/coccolith [or at least I think those are their units!). This is an overall range in PIC per coccolith for E. huxleyi of over 300X! Using different PIC per coccolith values in their model could obviously affect the PIC concentration mis-match, too! The authors need to state why the calcite per coccolith values are so variable and how much of the magnitude of the mismatch is due to what value of the PIC per coccolith is used. (The 300X range in PIC per coccolith makes this reviewer suspect that there must be some mistake in the units of the calcite per coccolith being discussed in this paragraph).
Thanks to the reviewer for spotting this mistake that we overlooked. It was indeed a typo in line 349: "4.64 pmol per coccolith" should be "4.64 pg per coccolith". We changed that and revised all the coccolith masses again to make sure the right units were used. We found a few typos that have now been fixed.

We also provided more details in section 3.2 (equation 1) regarding units.

In Figure 13 from Saavedra-Pellitero et al. (2019) we show this methodological difference in mass estimates, which is not 300X, but less than 1 degree of magnitude, but above all shows the same trend (this is also available in Figure 9 from this manuscript).

Due to the fact that some authors use pg and some pmol, we converted some of the published mass estimates to pmol (using the molecular weight of the calcite, i.e. 100 g/mol -section 3.2-) in the new Table 5 (which includes also new studies) and tried to stick to pmol in the discussion part to make it easier for the reader. Still mass in pg is sometimes mentioned if we refer to published papers, or in some of the figures (such as Fig. 9, which includes data from Saavedra-Pellitero et al., 2019).

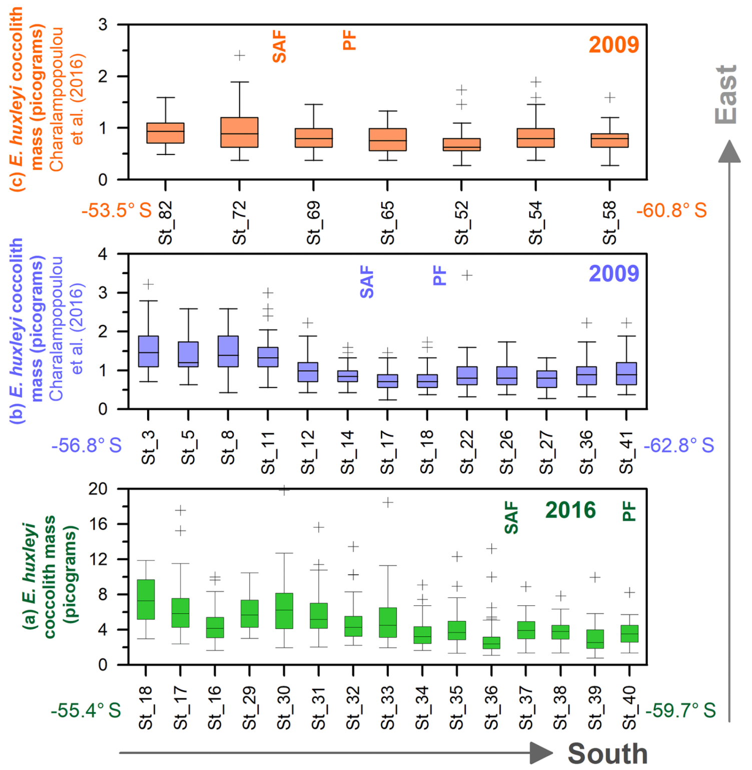

Figure 13 from SaavedraPellitero et al. (2019). Drake Passage latitudinal transects from east to west, showing box plot coccolith mass estimates (in pg): (a) this study, (b) transect at around 68◦ W from Charalampopoulou et al. (2016), and (c) transect at around 55–58◦ W from Charalampopoulou et al. (2016). Note that (b, c) have been calculated from Charalampopoulou et al. (2016). Outliers have been indicated with "+", and numbers on the x axes refer to the original station numbers. The approximate location of the Subantarctic Front (SAF) and Polar Front (PF) are shown as well as the year of the sampling.

They present discussion of the different morphotypes of E. huxleyi in this paper (lines 450-510). This reviewer is confused whether they are trying to connect the magnitude of the mismatch with the different morphotypes? The discussion seems more about simply presenting information about the different morphotypes that they encountered in the Southern Ocean. While interesting ecologically, does this have anything to do with the mis-match between SEM-based measurements of PIC and satellite measurements of PIC, the entire purpose of the paper? They could help this section by stating explicitly why they are presenting it at all. Only until I arrived at the very end of the paper (line 536) did I see conclusion #4 stating "neither the slightly different carbonate masses nor the southward changes in morphotype composition had a decisive influence on the coccolith-estimated PIC, which is only determined by the abundance of E. huxleyi in this area". This should be stated earlier in the discussion on morphotypes to tell the reader why you are including this morphometric data in the paper.

We acknowledge this point raised by the reviewer. We tried to convey it by adding some information already in the Introduction (mentioning the morphological diversity of *E. huxleyi*), in the methods (see new table 2) and later on in the discussion, with the effect of *E.huxleyi* morphotypes on PIC (see new Section 5.1).

Note that we deleted most of the previous section "5.3 *Emiliania huxleyi* morphotypes" in order to focus more on the PIC and less on the ecology of the different morphotypes.

Specific comments:

L 1-2 See general discussion above regarding making the title more descriptive of the paper.
We modified the title.

L 18-20 This sentence ("here we combine...") and the sentence in lines 21-23 ("We compare PIC estimates derived from...") are somewhat redundant. I'd eliminate the first.
Done; we deleted that sentence.

L 29 Change the words "coccolithophore data" to "coccolithophore-morphometric-based data"
We changed it to "coccolith-based PIC"

L34 The word "zooplankton" comes out of nowhere. Either eliminate altogether or state more carefully that you are referring to "calcifying microzooplankton."
We deleted the word zooplankton

L 41 Change "This process" to "Calcification"
We changed it.

L 54 Change "of the light backscattered..." to "of the total light backscattered.
We changed it.

L. 73 Understand the biological response to what?
We reworded this sentence as: "Recent concerns about climate change have motivated the scientific community to focus on *E. huxleyi* as a target cosmopolitan species and in particular to differentiate it into different morphotypes"

L 94 Change "different calcification levels" to "degrees of calcification of coccolithophore cells".
We changed it.

L 102 "Here, we focus on the..."
We corrected that typo

L105 Change "outer bloom conditions" to "moderate bloom conditions"
We changed it.

L111-114 sentence beginning "The ACC..." This sentence is long-winded. Please divide it in two.

We split this sentence into two shorter ones.

L149 The model of SEM is a Tescan Vega, I believe.
We changed it to “SEM Tescan Vega”.

L202 exchange word “reproducibility” for “precision”. Can you say anything about accuracy?
We changed it.
We understand that accuracy measures how close results are to the known or published value. We compare the available measurements in the discussion part.

L223 Refer to coccolith-estimated PIC concentration not just PIC
We changed it (and used satellite-derived concentration in that sentence).

L233 A zero SD could also mean there is only one measurement, right?
Yes, that is correct, but we feel there is no real need to mention this in the manuscript.

L239-242 Don’t refer to general “data scarcity” but refer to the more specific “lack of cloud-free satellite images”. Also, don’t say “to increase data availability” but “to increase the possibility of a ship-satellite match-up”
We changed those sentences.

L 239-249 See general comments on how they mixed level-2 daily satellite data vs level 3, 8-day and monthly Level 3 satellite data. Were Level 2 and Level 3 data used separately in different analyses or were they mixed in the same analysis?
See previous reply. We did not mix data from different levels. We used them independently. We reworded some sentences of this section to avoid confusion.

L 339 Change “calcite weight” to “calcite mass per coccolith”
We changed it.

L341 State what the PIC per coccolith numbers were used by Poulton, as well as Rigual Hernandez.
We added that information and rewrote part of that paragraph aiming to be more precise.

L348 Change “we calculated PIC” to “we calculated total PIC concentration”
We changed it.

L379- As the handling editor mentioned, stay away from the word “validation” in this paper because you are not doing this!
We agree with the Editor’s suggestion, but this refers to the validation by the NASA Ocean Biology Processing Group (2023) and not by us, so we kept it.

L383 Too many uses of the word “estimated” in this sentence
We swapped estimating for calculating.

L394 There is a need for both improved precision and accuracy!
We reworded this sentence

L397- Get rid of sentence on validation of remote sensing data since this paper is not about that!
We deleted that sentence

L435- You refer to “subfossil diatoms”. Are these different from fossil diatoms. Why not say “fossil diatoms in surface sediments”
Good suggestion, we changed it.

Line 547- Instead of “zooplankton”, do you mean calcifying micro-zooplankton like forams? Hard to imagine that your results will tell us anything about Calanus!
We changed it and elsewhere in the manuscript

Second revision of: Pacific Southern Ocean coccolithophore-estimated particulate inorganic carbon (PIC) versus satellite-derived PIC measurements

The authors have made some important improvements to the manuscript. However, I think it still needs some revision. I think they have addressed the technical issues brought up in the first revision. Reviewer 1 in the first round had very tough but also very valuable criticisms based on the emphasis on comparing satellite PIC to coccolithophore PIC without the needed chemical PIC measurements to do a true validation. I think they have gone far in addressing these concerns. However, they still focus on comparing satellite PIC vs microscopic estimates of coccolith PIC, when I don't think that's a strong comparison (as they lack chemical measurements of total PIC as pointed out by both reviewers).
We appreciate the reviewer's comments and would like to clarify a couple of points. We are aware that unfortunately we do not have in-situ PIC concentration measurements, and that the comparison is made using carbonate in coccoliths, which has its own inherent limitations. Additionally, the satellite data was often compromised by cloud cover, which further impacted the precision of our comparisons.

With that in mind, we consider that the value of our work lies in its ability to highlight trends and patterns despite these limitations. Our goal has never been to produce a 'proper' match-up with PIC measurements, but rather to provide a qualitative, comparative analysis that can offer insights into broader trends. We feel that this approach is still valuable in advancing understanding of satellite-derived PIC estimates, even without the ideal dataset for a true validation.

Given the constraints of our study, we hope the reviewer can appreciate that our work represents a step towards further research in this area. While it may not be a perfect comparison, we do believe it provides information of interest for the scientific community, and we have always framed our work as exploratory in nature.

We also would like to emphasise that all of these limitations – including the absence of in-situ PIC measurements, the reliance on carbonate in coccoliths, and the impact of cloud cover on satellite data – are clearly highlighted in the manuscript. We are fully transparent about these constraints and it is not our intention to present the results as a comprehensive comparison or to overlook these factors.

Again, I think the much more robust message, which also is more interesting, is that the high signal of PIC seen by satellite south of the polar front is clearly not due to coccolithophores. That is, their data help to define the true polar limits of coccolithophores. They could even present more clearly the sparse coccolithophores they find in those waters. The southern limit is an important observation, and I would say it is the main reason I think the work could be published.
It is true that our data can help to define the limits of coccolithophores in the Antarctic realm (even though we dealt with plankton samples -i.e. time snapshot-), so we added a few sentences in the abstract, discussion and conclusions regarding the southernmost extent of coccolithophores.

Still, we consider this is not the main aim of the paper. We explored the coccolithophore species ecologies in Saavedra-Pellitero et al. (2019) as well as Malinverno et al. (2015), and here we wanted to focus on a different (but related) aspect: the PIC.

We tried to make this clearer in the new version of this manuscript and we even shortened the discussion by deleting most of the section "5.3 *Emiliania huxleyi* morphotypes"

The main concern then is that the manuscript would just be essentially a re-publishing of previous results from the papers cited in Table 1. In that sense, I do think the comparison to satellite work does help in justifying publication, by adding to a new synthesis on defining how we can detect the southern coccolithophore limits. They may consider incorporating data published by other groups (some of which they cite) to support.
Only coccolithophore assemblage data was previously published. All the morphometric measurements and PIC estimates (on top of the satellite data) are new.
The suggestion of compiling published data from other authors to have a larger spatio-temporal span would be a great idea for future work.

The refocus would not justify such a long manuscript, and would need some condensing. Also, the manuscript requires some re-wording of the Title and Abstract, a modest reorganization of the Intro, and a more important reorganization of the Discussion to focus on that point.
Title, abstract and introduction have been modified, and notably shortened.

We reorganised some parts of the results and discussion; both sections are now much more concise.

In summary, I do support eventual publication, but I still think some further revision is needed. I evaluate these as "minor revisions" because I do not think a lot of effort or time would be required, though in fact it would still mean a major change of emphasis.
We hope we were able to provide it with the changes made.

Detailed comments:
Title: Still a problem with the title because of the word "versus"
We changed the tile as suggested also by the other reviewer

Abstract, lines 18-19: Again, remove focus on comparison "Here, we combine micropalaeontology and remote sensing to evaluate discrepancies between coccolithophore and satellite-derived PIC in the Pacific SO (in non-bloom conditions)."
This sentence has been removed as suggested also by the other reviewer

Lines 41-43: "This process decreases the alkalinity of surface waters, thereby reducing the uptake of CO2 from the atmosphere into the surface ocean and thus acting in opposition to carbon sequestration by the biological carbon pump (Rost and Riebesell, 2004)." The role of coccolithohores in ocean C uptake is probably more complex. For example, I am not aware that the ballast hypothesis has been completely discarded.
We modified that sentence and added the following one (and a new reference):
"Furthermore, coccolithophores influence the export of PIC and POC to the deep ocean

through the ballasting effects of their coccoliths into the deep sea (e.g. Klaas and Archer, 2002)."

Suggest moving much of the content of the second paragraph of the Intro to much later, so the paper has a focus on better understanding of the "Great Calcite Belt" and coccolithophore penetration of the Southern Ocean, rather than comparing satellite-PIC with coccoltihophores.
We notably modified the Introduction, and brought up the "Great Calcite Belt" earlier in this section.

Lines 106-107: "Our aims are: (1) to estimate the contribution of different coccolithophore taxa and morphotypes to PIC and (2) to compare coccolith-based PIC estimates with satellite-derived PIC values in the Pacific SO." I strongly suggest re-wording aims to match what can actually be achieved with their data. They cannot estimate the contribution of coccolithophores to total PIC without having chemical measurements of total PIC and resolving technical issues with estimating coccolith PIC by microscopy. So these aims must be completely re-written.
Boths aims have been reworded. We hope they are now clear and concise.

Line 125: "Coccolithophores dominate the SO phytoplankton communities" Despite the large number of references cited, I am not comfortable with the word "dominate", especially when considering the entire SO. They are important components of some SO phytoplankton communities, but clearly many SO phytoplankton communities are dominated by diatoms or by other phytoplankton (e.g. Phaeocystis)
We agree with the reviewer, and we reworded that sentence as: "Coccolithophores are important components of some of the SO phytoplankton communities".

Lines 133-139: "and almost always occurs as B morphogroup (types B/C and O). Furthermore, a general southwards decreasing trend in E. huxleyi mass, linked to a latitudinal trend from more calcified E. huxleyi (A morphogroup) to weakly calcified morphotypes (B morphogroup), was already recorded across the Drake Passage (Saavedra-Pellitero et al., 2019)." The concepts and importance of morphogroup and "more calcified"vs "weakly calcified" has not been introduced at this point, so appear out of nowhere. The importance of morphotype should be introduced in a re-organized and condensed introduction.
We added information regarding morphotypes and calcification in the introduction. This implied also adding a couple of new references.

183-184: "The importance of own size measurements for the determination of species-dependent coccolith PIC has been clearly emphasized (Baumann, 2004)". The phrase "own size measurement" does not seem to be correct English. Do the authors mean that it is important to make size measurements on the communities analyzed, rather than assume size measures from the literature?
Yes, we modified this sentence with the suggestion provided by the reviewer

Lines 203-206: "Coccolith volume estimates are likely to contain errors around 40-50% according to Young and Ziveri (2000), so we assumed the largest potential error and added a 50% error bars to our plots, although we note that measuring the actual size range in the

sample can reduce this error to about 5-10% in length and 15-30% in authors' attempt to respond to R1's detailed comments about needing better error analysis. However, I would prefer that the actual error bars refer to the variation in measure perhaps be indicated on the plot perhaps by lines or highlighting (with grey or color) what would be the expected range including this type of error (perhaps showing 15% and briefly how this uncertainty impacts conclusions (focusing on conclusions robust to the uncertainty).
This is indeed a good suggestion. We added a 15% error to (a) in Figures 3 and 4 and checked that both shaded areas (15% and 50%) are colour-blind friendly using https://www.color-blindness.com/coblis-color-blindness-simulator/
We also refer to this in the discussion.

Line 258- Subsection title: "4.1 Coccolith-estimated PIC versus satellite-derived PIC"
I strongly recommend moving away from a comparison of one estimation to another in a "versus" sense. Focus on defining the southern boundary of higher E. huxleyi abundances, and note that the satellite estimated PIC measures
We changed the title to: "Coccolith-estimated PIC and satellite-derived PIC" (which is now Section 4.2). As this is something already suggested by Malinverno et al. (2015) and Saavedra-Pellitero et al. (2019), we mention it in the discussion.

Fig. 3-4. I find it difficult to compare parameters with each parameter in a different panel with an offset. I think the figures could be re-organized. For example, one overlay panel (perhaps with right and left y axes) would be total coccolithophores, total detached coccoliths. Another would be E. huxleyi and C. leptoporos abundances. Another would be a column chart showing the relative abundances of the different E. huxleyi morphotypes.
We decided to update Figures 3 and 4 with other suggestions provided by this reviewer, but we did not follow his advice on this specific one.

We wanted to keep the coccolithophore species numbers (visually) separate from the morphometrics, and satellite parameters, so overlaying panels would mean a similar format to all the data. It is kind of a personal preference, but that is the main reason why we kept that part as it was before.

Fig. 5S should be mentioned in the first paragraph of section 4.1. It could even be a main figure after Fig.s 3-4.
We moved that figure from the supplementary material to the manuscript as a main figure (now Figure 7) and mentioned it in section 4.2 (we reorganised the results part, so section 4.1 is now 4.2), but in the second paragraph of that section.

Fig. 2 and Fig. 6. How is tube width measured in the Type A overcalcified coccoliths where the tube overgrows the central area in an irregular manner? How is T element width overcalcified coccoliths where the elements are fused? Where is the data on element width?
This is a very good question. It was very challenging to measure the tube width when it was overgrown in an irregular manner. We aimed for the best possible fitting in the manual measurements (New Zealand samples), and we trusted the Cocobiom 2 macro to automatically do the same.

T elements were measured, when they were fused or it was not possible to measure them for any reason, we indicated it with N/A (not available). Still that data, even if it is available in tables 1S and 2S in the supplementary material.

Fig. 7 and Fig. 8 should be presented in Results.
Previous figure 8 is mentioned in the Results section, but previous 7 (now 9) is only mentioned in the Discussion section.

The Discussion should start with defining the southern boundary, based on the new results here as well as reviewing previously published microscopic evidence.
We re-arranged the results and discussion part.
We deleted the previous section 5.3. and kept just two sections (which we rewrote) in the discussion part to make it more concise and focus on the main topic: the PIC estimates.

Note that several references were deleted while rewriting the introduction and discussion parts.