*We kindly thank Dr Robert Spielhagen and Dr Ralf Schiebel for the constructive feedback on our manuscript. We are thankful for the remarks on the content and proposed changes to improve the quality of the analysis.*
*Our detailed responses to all raised issues and the changed made to the manuscript can be found below, written in italics below each of the comment made by the reviewers.*

**Response to RC1:** 'Comment on bg-2022-59', Robert F. Spielhagen, 23 May 2022

General: Be consistent with the use of either British or American spelling (BE/AE): -ise/-ize, -our/-or, etc.
*We have checked the whole manuscript and changed it to be consistent with British spelling.*

Fig. 1: The map should be concentrated on the actual sampling areas. A cut-out square concentrating on the Arctic Ocean, the Nordic Seas and the Labrador Sea/Baffin Bay area will show these areas in higher resolution.
*Following this advice, we changed the map in the revised version such that it only shows the areas relevant for the study.*

Line 162: For a better guidance of the reader you may write "in 37 out of X profiles" (where X is the total number). I am aware that the total number of profiles was mentioned earlier, but it is helpful if you repeat it here.
*We changed the sentence to "This was the case in 37 out of 126 profiles".*

Lines 164, 169: the BPZ
*We changed this.*

Line 217ff: Consider giving 1-sigma or 2-sigma values for some important data (e.g., mean depths of certain parameters) also in the text. Calculated averages (means) are important, but many readers may be also interested in, e.g., the depth range of BPZ values (and others).
*To not diminish the readability of the text, we have incorporated information on the variance not in the text but in Table 4. As the sampling sizes vary substantially among the analyses, we present 95 % confidence intervals.*

The scales in your figures often have only few (or no) minor ticks between labeled ticks (e.g., water depths) and it is hard to visually determine numbers between these labeled ticks.
*We have increased the number of axis ticks and labels* in all figures where there were only few ticks present.

Fig. 3a: Needs additional explanation in the figure caption on what is shown here (for those who are not experts on statistical methods;-). What is the box? What is the vertical bar in the box? What is the horizonal bar? What are the black dots? If you give this information here, you can refer to it in other figure captions.
*To make the content of the figure clearer, the following explanation was be added to the figure caption: "The box represents the interquartile range (IQR) of the relative abundance at the given depth, and the vertical bar represents the median. Outliers, shown as points, are values beyond 1.5*IQR of each site of the box, and lines represent the range within 1.5*IQR."*
*We also refer to this explanation in the captions of the figures 4, 6 and 9, which show analogous plots.*

Table 2 and lines 228, 248, etc.: Expedition numbers, station numbers and deployment numbers from Polarstern expeditions should be used correctly in the manuscript and not be mixed. Even if individual sample containers or bottles may have been labeled differently

during the expedition, links to the PANGAEA data bank (and to, e.g., CTD data from the same station) will work properly only if the authors make a correct use of the numbers which are different in certain details. In the case of the Polarstern expedition in 2015, the following scheme applies:
PS93.1 is the expedition number (i.e., the first leg of expedition PS93 which was split into two legs during a late part of the planning stage).
PS93/20 (or PS93/24 or similar) is a station number (NOT PS93.1 20!).
PS93/20-3 is the deployment number of the multinet haul at station PS93/20 (PS93/20-1 was a handnet deployment, PS93/20-2 was a CTD run, etc.).
A correct use of these numbers is essential in many ways. A similar scheme is applied to expeditions of other vessels (e.g., Maria S. Merian, Meteor etc.).
*We thank the referee for this important comment. It is indeed essential that the labelling makes the data unambiguously identifiable and reusable. We therefore checked all labelling within the manuscript and changed it to be in accordance with labels used in PANGAEA.*

Figs. 5, 7, 10, 11: It may be related to my poor experience with logarithmic data presentation, but I find the labeling of scales a bit confusing... For example, I assume that the horizontal scale in Fig. 10 shows log[shell flux], given in mg per square meter per day(?). Wouldn't it be possible to label the scale ticks as 0.01, 0.1, 1, 10, 100, 1000 ? Similar changes can be made to the other figures with logarithmic scales. Add minor ticks between the labeled ticks, because our brain is trained to think on linear scales, not logarithmic ones…
*We have changed the labels of scales of all the named figures to show absolute values but plotted on a logarithmic scale to facilitate interpretation.*

Lines 375/376: This statement needs at least a reference to a figure (or some explanation). The given number is somewhat confusing because in chapter 3.1 you give 124 and 136 m as BPZ depths. Different ways of calculation? Which number should be cited if future authors want to use your work as a reference? You may also consider rephrasing the sentence ("The average BPZ, calculated as ..., is at ...).
*We have change the sentence to show the difference between the overall average value of the calculated BPZ in contrast to BPZ averages based on the different calculation methods explained in the methods section as follows: "*Our analysis of observations from plankton net samples indicates that the productive zone of *N. pachyderma* in the Arctic and Subarctic realm reaches down to about 113 m water depth (median of all samples**, 125 m for where calculation after Lončarić et al. (2006) was possible, 136 m defining it as the range end**)."

Lines 437/438: "... rules out that empty shells in the upper water column only represent specimens affected by premature depth." - I do not fully understand what you want to say here... What do you consider the "premature depth"?

*We are sorry for the wrong spelling, which should be "premature death". We changed this and further added an explanation in brackets to the exact meaning here: (foraminifera which died at a juvenile or young adult stage, likely without reproduction)*

Line 444: Upper or lower end?
*We changed the sentence as follows to make clear that the lower end is meant here: "Ten out of 18 of the here studied profiles indicate an increase in calcification intensity **with increasing depth** within the productive zone".*

Lines 454-456: Horizontal advection may also play a role, for example in the narrow Fram Strait where specimens from the east (with a thick Atlantic Water layer) may be transported into strongly stratified waters with a top layer of Polar Water.

*We included a new sentence in the manuscript referring to v. Gyldenfeldt et al. (2000) who demonstrated that N. pachyderma might also be transported, with the corresponding value in*

*the Fram Strait amounting to a transport distance mean of 25-50 km in the upper 1000 m for sinking foraminifera shells, which can play a role for the collected material: "Furthermore, sinking shells of N. pachyderma can be transported over considerable distances, as e.g. shown by v. Gyldenfeldt et al. (2000), whose results would indicate a transport of 25-50 km in the upper 1000 m, resulting in the possibility of some of the encountered specimens being advected from areas with a different hydrography."*

Line 476ff: For the reader who does not always want to jump back to the Results chapter, it will be useful to find some numbers here when data and data ranges are discussed. For example:
- 492: "... fluxes that are three to five times higher than estimates..."
- 497: "The highest estimated calcite fluxes in our data set..."
*We have included relevant values in brackets when comparison of numbers where done in the discussion.*

Line 534-540: If mesh size is so important, why don't you determine this factor also within your data set? I am aware that many other factors (water temperatures, summer ice coverage, nutrients, etc.) also influence foraminifer fluxes. However, within your large data set you should be able to compare results from studies using different mesh sizes within the same region (where conditions are roughly the same) and even do this for several regions. For example, if you have, say, five studies in the western Fram Strait and western Greenland Sea using 150 microns and another five with 63 microns, you should at least be able to get an idea of the effect of different mesh sizes. This may statistically not be absolutely significant, but it may help to understand the effect.
*We are thankful for this suggestion for a useful addition to our study. We have checked the effect of different mesh sizes used for sampling on the abundances and added the results of it in the manuscript as follows: "A comparison of abundances of N. pachyderma in our compilation derived from the same region, but sampled with different mesh sizes, shows that its abundance is on average 27 % lower when a coarser mesh size (100 µm, 125 µm, 150 µm) is used, because small shells are not sampled. These observed estimates of a reduction in the abundances is comparable to the results by Carstens et al. (1997), who detected a reduction in foraminifera abundances of 7 % to 40 % with increasing mesh size. The flux given by Schiebel (2002) is based on data from sampling with a 100 µm mesh size. Our data from the western Fram Strait indicates that in this region, the abundance of larger (>125 µm, >150 µm) shells is on average 56 % lower than what is sampled with a mesh size of 100 µm. With 49 out of 148 stations in our dataset having a mesh size coarser than 100 µm, the lower flux estimates in our compilation are likely at least partly underestimated, compared to fluxes consistently based on sampling with a mesh size of 100 µm, but the difference is unlikely to be larger than one third."*

Line 543: Schiebel
*We changed this.*

Line 553: Insert commas before and after "taking the abundance of N. pachyderma into account".
*We changed this.*

Line 572: contribution of planktonic foraminifera
*We changed this.*

Line 572/573: and a higher
*We changed this.*

Line 572ff: Is it really justified to discuss "the Fram Strait" as an entity? Concerning environments, the western FS resembles large parts of the Greenland Sea, while the eastern FS is more similar to the northern Norwegian Sea (strong near-surface influence of Atlantic

Water, higher temperatures, no sea ice...). I do not think that one can say that there is a stronger influence of Atlantic Water in the Greenland Sea than in "the Fram Strait" (as stated in line 576). The advection of AW is largely meridional in the Nordic Seas, with parts of AW branching off to the west in various regions, including the Fram Strait. This (and the potential influence of mesh sizes) makes the statements in lines 576-578 rather vague...

*We agree that considering the Fram Strait as a homogenous region is not entirely correct. We therefore divided the Fram Strait and the Greenland Sea into western and eastern part in our analysis to enable a more precise comparison. We have changed the paragraph in the manuscript as follows: "A direct comparison of fluxes of planktonic foraminifera from samples from within the same region with total $CaCO_3$ fluxes in the region indicates a lower contribution of planktonic foraminifera in the Eastern (> 0°) Fram Strait (10 %) and a higher contribution in the western part (< 0° E) of the Greenland Sea (50 %) to total $CaCO_3$ fluxes. For this comparison, we subdivided the regions by longitude to account for the different influences of Atlantic and Arctic waters, which play an important role for the abundances and habitats of planktonic foraminifera in this region (Pados & Spielhagen, 2014). The contribution of 10 % in the Fram Strait is in line with the lower end of estimated contribution of planktonic foraminifera to total $CaCO_3$ fluxes at the Northern Svalbard margin (4-34 %; Anglada-Ortiz et al., 2021). The higher contribution in the Greenland Sea is in the range of the estimates from Salter et al. (2014) from the Crozet Plateau in the Southern Indian Ocean, indicating that the given contribution falls within globally realistic ranges. The previously described possible effect of coarser mesh size decreasing flux estimates has to be considered, meaning that the values from our dataset provide a minimum range."*

Lines 579/580: ... similar ... as ……
*We changed this.*

Line 581: than in other
*We changed this.*

Line 606: Arctic Ocean
*We changed this.*

Reply to **RC2**: 'Comment on bg-2022-59', Ralf Schiebel, 13 Jun 2022

Overall, the manuscript on by Tell and coauthors reads goods, and adds useful information on the population dynamics and flux of N. pachyderma. In addition to the comments and suggestion made by another reviewer, Robert Spielhagen, I would suggest amendments, which I have detailed in an annotated pdf file I send along with this review. In particular, wording needs to be revised in places. Some of the figures are not up to publication standard, and labels need to be revised and/or added. Most importantly, the manuscript needs to be rearranged in places, and some sections need to be moved from the "Results" to the "Discussion". I would also suggest to analyze the data on test size for any relationship with the synodic lunar cycle, which would strongly support (or not) the population dynamics and pulsed flux discussed in the manuscript.

*We have incorporated all comments on the manuscript regarding grammatical issues and typos. We have also specified the wording and changed "calcite" to "carbonate" where necessary and "production flux" to "export flux" in the whole manuscript. Where pointed out, we have added further literature references.*

*We are thankful for the suggestion to analyse the possible linkage between foraminifera test size and the lunar cycle. We have tested the relationship between shell size and lunar day*

*using the size throughout all depth intervals, divided into the different time intervals and regions of sampling, represented by the three different cruises where size data have been generated. In this analysis, we did not find any evidence for a linkage between the lunar periodicity and the shell size of N. pachyderma. But our data is also not providing enough material to analyse this in detail, as the size data we have does not span over a whole lunar cycle. We have added one sentence on this in the result section 3.2 ("We find no evidence for lunar periodicity in the shell size of N. pachyderma in our samples.") and further specification in the discussion section 4.2 ("Our data also does not present a strong systematic change in size with lunar day, as it was detected in previous studies (Schiebel et al., 2017). However, our shell size data do not cover the entire lunar cycle, preventing drawing firm conclusions on the influence of the lunar cycle on the shell size of N. pachyderma.")*

L. 7: that's a very anthropocentric point of view, and varies between species
*We have changed it to "could become less challenging to pelagic calcifiers".*

L. 44: and biological carbon pump
*We added this.*

L. 52: deeper than what?
*We added "than the productive zone", to clarify what "deeper" refers to.*

L. 53: hm, maybe or maybe not; this is not a scientific argument, and I would suggest to skip it
*We changed the sentence accordingly: "Also, sediment trap records are too scarce in the Arctic..."*

L. 68: and? ... for the sedimentary assemblage while non-encrusted tests have been removed by dissolution...?
*We have changed the second half of the sentence as follows to make this more precise: "because most individuals add a crust and because encrusted shells are more resistant to dissolution*

L. 72: How do you define a "crust"? Gametogenetic calcification?
*We have changed the following explanation in the methods section: "Upon sampling, no direct differentiation between shells with or without a crust was done. We regard shells as having built a crust based on their larger weight, different shell texture and more rounded shape, as exemplarily shown in Fig. 2g."*

L. 86: an individual?
*We have changed the sentence accordingly to make this clear: "To distinguish the production and export zones and to determine the average depth of calcification of N. pachyderma…"*

L. 104: processed two times? processed for what?
*As this was unclear in the manuscript before, we changed the sentence: "Samples from the Baffin Bay were either processed on board or stored at -80°C until processed onshore."*

L. 106: usually, the minimum diameter is assessed, because it is comparable to sieve-size intervals
*We unfortunately only have the measurement of the maximum diameter for the samples from the Baffin Bay. We agree that using the maximum diameter does not enable a direct comparison of absolute values to sieve-size intervals. The most important aspect, nevertheless, is to not mix different types of size parameters in one analysis, which we do not do. We are therefore confident that our conclusions are valid. We therefore stick with the*

*parameter as done in the before submitted version of the manuscript to be able to work with the datasets together.*

L. 111-112 / Fig. 2 (f): To my impression, the specimen shown in (f) has not been alive when sampled, and Rose Bengal staining may not be suited for live-staining of planktic foraminifers.
Please see the paper of Lutze and Altenbach (1991) Technik und Signifikanz der Lebendfärbung benthischer Foraminiferen mit Bengalrot.

*We are aware of the fact that rose Bengal might be staining recently dead specimens because of cell degradation (see Schönfeld et al., 2014). But visually separating shells based on the presence or absence of cytoplasm without staining also leads to some ambiguity. This is because recently dead foraminifera can still contain cytoplasm. Therefore, the proportion of dead foraminifera represents in most cases lower bound/slight underestimation. This also becomes clear by our results shown in the manuscript in Fig. 3a. We included remarks on this in the method section:*
*"The counts were made separately for cytoplasm-bearing shells and empty shells, differentiated during the processing of the wet samples. As recently dead foraminifera can still contain cytoplasm, this leads to a bias in the numbers in favour of shells interpreted as being alive upon sampling."*
*"In accordance with data from earlier studies, fully white shells were classified as dead ("empty") (e.g. Fig. 2e), all other (pink) shells as cytoplasm-bearing (e.g. Fig. 2f), assumed to represent specimens that were alive during retrieval. As rose Bengal might be staining recently dead specimens because of remaining cytoplasm in the shells (Schönfeld et al., 2014), there is a possible bias towards too high numbers of cytoplasm-bearing shells."*

L. 160: 37 profile is a lot, and a different method than Loncaric's may be applied for a more representative result
*It is a valid point that the method has disadvantages. We nevertheless suggest keeping it, as we think that our overall estimates on the BPZ are satisfying enough to see what changes are present within the productive zone and where it ends to be able to calculate carbonate fluxes below. We added a sentence in the manuscript to make the disadvantages clear: "For those 40 profiles, the BPZ was defined as the bottom depth of the transition zone (Fig. 2a, $Z_{BPZ}$ (range end)). This can result in a bias towards the estimated BPZ being located below the actual position. This bias is restricted by the overall sampling interval (median: 50 m) and has no effect on our flux estimates which are based on average shell abundances below the BPZ."*

L. 180 ff: this section reads like a Discussion, not Materials and Methods
*The intention of this section is to explain why we investigated the named parameter and what aspects we looked for in our analysis, which is why we put it into the methods section. We think it is necessary to have it in this section to help the reader understand the purpose of the analysis. Therefore, we would like to keep it at its current position. After naming the three parameters, we have added "The reason for using those parameters is that if…" to make clear that this is not discussion, but an explanation on why analysing the named parameter.*

L. 198/199: this may not possibly give realistic results, because the population dynamics and flux is too dynamic and too fast
*We agree and are aware of the fact that using average values could be rather far from reflecting the actual situation. Nevertheless, as long as no further measurements are available, we think this is the simplest way to estimate mass fluxes in different regions of the Arctic realm. To make clear that this can have an effect on the analysis, we have added a sentence to the manuscript: "This method is likely to underestimate present variability."*

L. 220: subsurface is possibly not defined by depth alone, but by surface water stratification, i.e., subsurface under the seasonal pycnocline

*We agree, but we are here not referring to a defined subsurface in terms of the water masses. We just intend to express that some profiles show a maximum directly at the 0-50 m depth interval, and some only further below (50-150 m depth). To avoid confusion, we have changed the sentence as follows: "Shell abundances show either a maximum within the upper 50 m, or in the depth zone below, reaching down to 150 m (exemplarily shown in Fig. A1)."*

L. 274: is this rounding, or is it 0.010 ? Please use the same number of digits for the same purpose
*Thank you, it is 0.010, we changed this.*

L. 281 do you mean "less lobate"?
*Yes, we changed this in the text.*

L. 285 in "roundness"?
*Yes; we added "in the roundness of shells" to the text to make this clear.*

Fig.7: change labels of both x- and y-axis to meaningful information
*We would like to keep the label of the y-axis as it is, because another label could be even more confusing. However, we understand the point and therefore added specification in the caption of the figure: "100 % equals the total depth of the productive zone, 50 % half of the depth of the productive zone. More than 100 % are reached where the sampling interval ends below the BPZ." We also changed the labels of the x-axis values to absolute values so they can be interpreted directly.*

L. 315 in what?
*A decrease in mass flux, we specified this in the text.*

L. 317: what do these numbers show?
*To make clear that the numbers in the brackets in this paragraph always represent the same, we have specified it more clearly in the first brackets: "20.1 mg $CaCO_3$ $m^{-2}$ $d^{-1}$ based on weights of encrusted/empty shells only; 4.5 mg $CaCO_3$ $m^{-2}$ $d^{-1}$ based on weights of non-encrusted/filled shells only; in the following, those two values will always be given in brackets without further stating this specification".*

L. 320: this not "Results" but "Discussion"
*We deleted this part here, and also do not bring it up like this in the discussion, as re-calculation of the values changed the here stated effect.*

Fig. 8: What does "Count" include? Any unit? Please add numbers to the "open ends" of the x- and y-axis. The "floating bar" at the very left end of the x-axis looks strange. Please think about re-designing the figure.
*We chaned the label to "Number of observations" (it is not the count of foraminifera, but sampled stations) and increased the number of axis ticks in the figure to improve its quality.*

Fig. 9:
"Change" of what?
What are the scales and units of the dots and boxes in the top panel?
is this the orange color?
*We have changed the labels of both plots to "change in mass flux [%/100 m] to make clear what is shown here. We have deleted the legend in the figure and use the figure caption to explain what the differently coloured dots represent. The named colours have been changed to "blue" and "orange". We further added an explanation on what the box plots on top of the figure represent: "The boxes and bars on top of the plots represent the interquartile range as explained in the caption of Fig. 3 and are plotted against the same x-axis as the plot below."*

Fig 10: what is the width of the intervals?
*We have clarified this in the figure captions in the new version of the manuscript: "The exact width of sampling intervals differs between individual sampling locations. Details on this are shown in Tbl. 2 for profiles added in the study, and in the linked references listed in Table 1."*

Fig 12: how realistic are residence times of >80 days?
*It is not possible for us to exactly predict what residence times are really realistic for N. pachyderma in the Arctic realm. Based on culture experimentations (Spindler, 1996), however, we know that the genotype from the Southern Ocean can live several months and up to more than a year. It is equally not implausible that the extremely long implied residence times are a product of partly unprecise flux estimations. We added explanation on this in the discussion section to make this clear: "Nevertheless, the estimated residence time of about 4 days in our data indicated that the life span of the sampled N. pachyderma is either too short to be strongly affected by environmental condition changes, or that the population size is constant at least across a short time scale. The latter would make huge changes in the environmental conditions unlikely. When interpreting the lower end of estimated residence time, it has to be considered that this might not represent the overall lifespan of the foraminifera, but the days it stays alive after having reached maturity, which is what, based on optical parameters, we mainly analyse. Based on results from culture experiments on N. pachyderma from the Southern Ocean, a lifespan of several months, the upper end of our estimated residence time, seems to be possible (Spindler, 1996). It is equally possible that unprecise flux estimations in some samples lead to those extreme values. With the majority of all samples showing a residence time of only a couple of days, we conclude that the possible blurring of signs of OVM would be rather small, and the lack of a clear trend indicating OVM at all stations can be seen as a reliable result."*

L. 376ff: please discuss also the results presented by Carstens, Wefer, Volkmann, Pados, Spielhagen, Schiebel, etc
*We can included the results from Carstens and Wefer in the discussion. As we use all the data from the other studies in our manuscript, it is trivial to make a comparison. We have added the following discussion in this section: "A variation in the depth interval of maximum abundances of N. pachyderma is also presented by Carstens & Wefer (1992) and Carstens et al. (1997), where a connection between distinct water masses and temperature regimes is drawn. Our dataset corroborates these observations and indicates that the base of the productive zone of N. pachyderma is also highly variable and reflects the habitat depth (vertical distribution of living specimens). Like Greco et al. (2019), we observe that even if there would be a general pattern of habitat depth and BPZ position being driven by environmental factors, as also proposed by Carstens et al. (1997), it is overlain by considerable variability, even among profiles collected in the same region and around the same time."*

L. 402 by dissolution?
*Yes, we specified "due to dissolution" in the new version of the manuscript.*

L. 404-405: i.e., statistically not significant? Please show in the "Results" section!
*We did not do statistical analysis exactly because of the low numbers. We therefore changed the wording to make clear that we here refer to* descriptive results only*: "A high percentage gain in flux might in some cases only represent a difference of a few shells, which is not related to an actual higher flux but to methodological uncertainties of sampling."*

L. 407-408: please present the statistical significance, and, if significant, discuss the results
*We did not do statistical analysis exactly because of the low numbers. We therefore changed the wording to make clear that we here refer to* descriptive results only*: However, we assume*

*that the uncertainties concern all data in the same way, they should not be systematically affecting individual profiles. Therefore, we conclude that the overall pattern of a variable BPZ often located at and even below 100 m, but never below 300 m, should be reliable."*

L. 409: These observations can help to calculate fluxes - please do
*We changed the sentence as follows: "These observations helped us to calculate fluxes of shells of N. pachyderma from plankton net samples at a more realistic depth in the Arctic Ocean, even where the productive zone has not been explicitly constrained. It can serve as a base for further studies."*

L. 413-425 what about the Schiebel et al. (2017) data on depth distribution and reproduction of N. pachyderma? Please discuss
*See our comment on our analysis regarding the lunar cycle at the beginning.*

L. 455-456: any proof?
*There is no proof, which is why we stayed speculative with our wording: "Depending on the life span of N. pachyderma, which could be longer than one or two months (Carstens and Wefer, 1992; Kohfeld et al., 1996), it is possible that the samples contain individuals from multiple generations that were produced during different environmental conditions."*

L. 457-458: do you suggest a life span of N. pachyderma of 4 days?
*We understand the confusion, see our addition to the manuscript to the discussion section we explained above in connection to the comments to Fig.12.*

L 483: what?
*It refers to the shell weight, we clarified this in the manuscript.*

L 483 ff please rephrase this paragraph to make it concise and unequivocally understandable
*We rephrased the sentence to make the effect of different mesh sizes during sampling clear: "The used sampling mesh size creates a bias, as a 63 µm net samples different material than a 100 µm or 125 µm net. This bias would result in higher average shell weights when a coarser mesh size is used for sampling."*

L. 495: sorry, I've got no idea what's meant here
*We rephrased this paragraph to make clear why we use the overall average shell weight and not only that of empty and crusted shells as follows: "However, our observations indicate that not all specimens build a thick crust before reproducing and dying and some still contain remainders of cytoplasm while already sinking. Therefore, flux calculations based on averages of all shell types should be more realistic then only using weights of encrusted, empty shells."*

L. 498: to calculate flux. (this should be movded to the "Results" chapter.
*We added a reference to what was explained in the method section here: "We do not have any weight measurements from the Baffin Bay, hence use overall averages to calculate fluxes for data from there, as explained in the method section (Sect. 2.3)."*

L. 542: not on average. However, following deep mixing, e.g., by storms, the surface mixed layer and productive zone of planktic foraminifers, may be deeper than 100 m, and which may have a huge effect in particular during times of enhanced test production in spring. Please see Schiebel et al. (1995).
*We agree, which is why we did not limit our flux calculation to 100 m. We changed the sentence accordingly to make clear that we cannot judge the effect of a different productive zone depth in the North Atlantic in terms of the comparison with this data set: "We show that 100 m can be too shallow to estimate the fluxes in the Arctic, but cannot judge the effect of a*

*possibly deeper or varying productive zone in the North Atlantic (Schiebel et al., 1995) on flux estimates"*

L. 595: this is not far away from the 100 m used as export layer by Schiebel (2002), isn't it ;-)
L. 598: if this is true, how do you arrive at a 5-fold overestimation when comparing your data to Schiebel's (2002)? I guess that you are may be right with your assumptions, but you should present it in the right way

*Thank you for pointing out that the current formulation in the concluding chapter is not clear enough to well present the main results of our study, and to make clear why the difference to the 100 m used by Schiebel (2002) is relevant, even though it is not far away from it. We have rephrased the paragraph as follows: "Our compilation of vertically resolved data on the dominant Arctic planktonic foraminifera N. pachyderma reveals that the base of the productive zone of this species is on median located at about 113 m depth, but shows large regional variability and locally reaches down to 300 m. Our analyses show that it is important to constrain the base of the productive zone to estimate fluxes in the export flux below: using a constant 100 m depth to estimate fluxes leads to a fivefold flux overestimation in contrast to the flux at the top of the export zone. Below the BPZ, the shell flux is decreasing on average by 6.6 % per 100 m, with highest losses directly below the BPZ. No further change is observed deeper than 300 m below the BPZ. Therefore, we can conclude that in the absence of knowledge on the position of the BPZ, using 300 m depth should provide a conservative, yet more realistic estimate of the N. pachyderma export flux in the Arctic realm than using the formerly often used depth of 100 m. Within the productive zone…"*