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

*Please find our detailed responses to all raised issues and our proposed changes to the manuscript below, written in italics.*

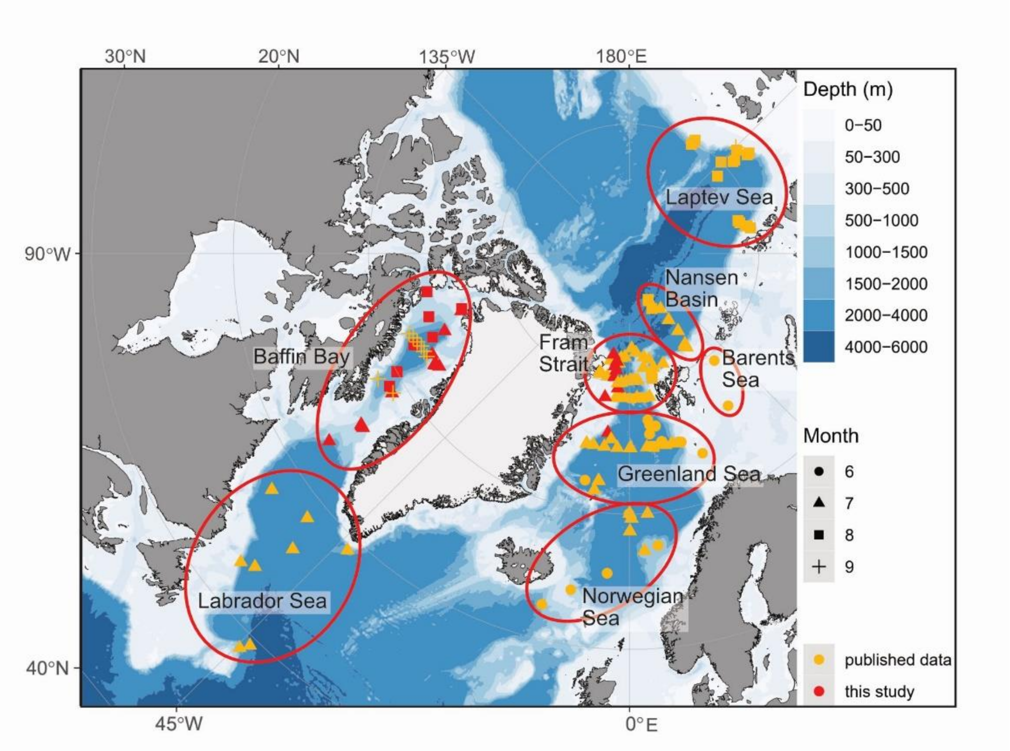
The manuscript by Tell et al. discusses published and new data sets of planktic foraminifers from plankton tows in the Nordic Seas, Labrador Sea/Baffin Bay and the adjacent Arctic Ocean concerning the quantification of calcite production by the dominant polar planktic foraminifer N. pachyderma. This is a species widely used by paleoceanographers for high-latitude paleoenvironmental reconstructions and any additional knowledge on the behavior of this species in the water column, the fluxes and potential burial rates is highly welcome. The manuscript provides this additional knowledge by offering a holistic view on an yet unprecedented number of data sets and is therefore of high significance for the scientific community. Since the manuscript is well written in excellent English, well organized, and equipped with illustative figures, it will make a valuable contribution and is well suited in "Biogeosciences". My limited knowledge of statistical methods did not allow me to evaluate all the results from such methods in detail and I may have overlooked some critical points. I hope that some more experienced experts can give some comments on these. However, my overall impression is very positive and I suggest publication in "Biogeosciences" after a minor revision which considers the specific points listed below.

General: Be consistent with the use of either British or American spelling (BE/AE): -ise/-ize, -our/-or, etc.

*We will check the whole manuscript and change 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 intend to change the map in the revised version such that it only shows the areas relevant for the study (see below).*



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 will incorporate this and agree that this is a helpful addition to table 3.*

Lines 164, 169: the BPZ

*We will change 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).

*Providing not only mean values but also the variance information would indeed be useful. However, we believe that adding this information in the text would make it less easy to read and we therefore propose to present the information in Table 4. Because the sampling sizes vary substantially among the analyses, we believe it will be more appropriate to present 95% confidence intervals instead.*
*Table 4: Overview on measurements on different shell parameters from the different sampling areas of the study. Next to each indicated value, the 95% confidence interval (CI) is given in italics. It is calculated assuming normal distribution, and the number of samples (n) used to calculate each parameter is given in brackets.*

| Shell abundance (ind. m-3) | | | | BPZ (m) | | Shell size (µm) | | | Shell weight (µg) | | | Mean mass flux (mg CaCO3 m-2 d-1) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mean productive zone** | *95% CI (n)* | **Mean export flux zone** | *95% CI (n)* | **Median** | *95% CI (n)* | **Mean** | **Median** | *95% CI (n)* | **Mean** | **Median** | *95% CI (n)* | **100m** | *95% CI (n)* | **Below BPZ** | *95% CI (n)* | **Deepest sampling position** | *95% CI (n)* |
| 42.3 | *6.1 (404)* | 6.5 | *0.0 (360)* | 112.7 | *10.8 (127)* | 150.0 | 143.4 | *0.5 (40275)* | 3.4 | 2.3 | *0.3 (330)* | 40.8 | *10.3 (173)* | 8.0 | *3.0 (126)* | 4.4 | *2.3 (126)* |
| 90.4 | *24.0 (71)* | 16.4 | *6.6 (77)* | 100.0 | *28.3 (22)* | 146.5 | 140.6 | *0.5 (36070* | | | | 85.5 | *26.1 (44)* | 22.7 | *14.3 (21)* | 13.7 | *12.8 (21)* |
| 0.6 | *0.4 (7)* | 0.5 | *0.3 (3)* | 160.3 | *51.8 (2)* | | | | | | | 0.7 | *0.6 (2)* | 0.5 | *0.5 (2)* | 0.5 | *0.5 (2)* |
| 19.5 | *3.6 (126)* | 3.2 | *0.9 (120)* | 100.0 | *16.7 (44)* | 180.0 | 172.1 | *1.6 (4205)* | 3.0 | 2.1 | *0.4 (193)* | 18.3 | *5.3 (50)* | 3.9 | *1.6 (44)* | 1.6 | *0.9 (44)* |
| 79.2 | *36.5 (56)* | 9.9 | *4.3 (45)* | 116.1 | *27.7 (14)* | | | | 3.0 | 3.0 | *0.5 (40)* | 66.6 | *45.4 (23)* | 17.1 | *10.6 (14)* | 8.0 | *4.1 (14)* |
| 14.2 | *11.5 (20)* | 2.6 | *1.5 (18)* | 153.0 | *59.7 (8)* | | | | 4.4 | 2.8 | *0.8 (97)* | 6.6 | *3.4 (8)* | 2.4 | *2.0 (8)* | 4.3 | *4.8 (8)* |
| 2.8 | *0.8 (41)* | 0.7 | *0.2 (35)* | 139.7 | *31.9 (15)* | | | | | | | 4.3 | *1.7 (16)* | 1.0 | *0.4 (15)* | 0.5 | *0.4 (15)* |
| 35.8 | *7.1 (26)* | 2.9 | *0.9 (24)* | 106.4 | *36.8 (10)* | | | | | | | 34.8 | *7.9 (10)* | 5.1 | *1.2 (10)* | 2.3 | *1.6 (10)* |
| 42.7 | *23.2 (57)* | 2.8 | *1.0 (38)* | 125.8 | *51.4 (12)* | | | | | | | 19.1 | *8.8 (20)* | 3.0 | *2.5 (12)* | 1.8 | *1.3 (12)* |

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.

*Indeed, such change would improve the presentation of the results and we propose to add more axis ticks and labels in all plots that have large distances between the given labelled ticks. Please see the following figure as an example for the way we will change it in all figures.*

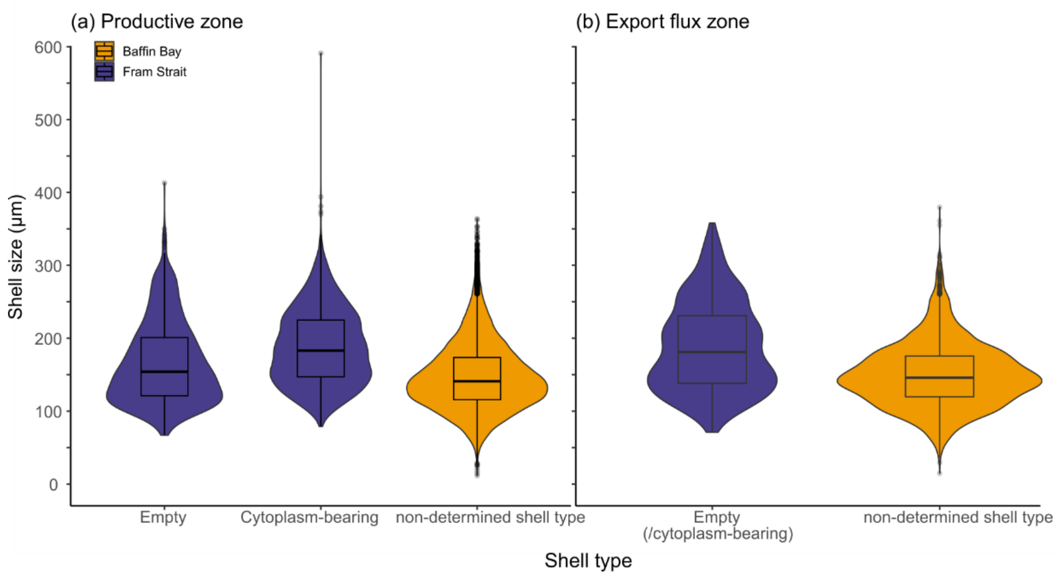


Fig. 3a: Needs additional explanation in the figure caption on what is shown here (for those who are not experts on statistical methods;-). What is the box? What is the vertical bar in the box? What is the horizonal bar? What are the black dots? If you give this information here, you can refer to it in other figure captions.

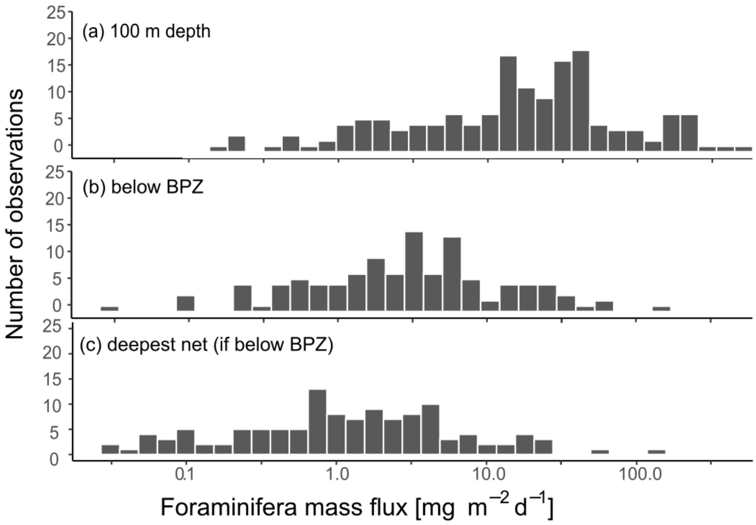
*To make the content of the figure clearer, the following explanation will be added to the figure caption:*
*"The box represents the interquartile range (IQR) of the relative abundance at the given depth, and the vertical bar represents the median. Outliers, shown as points, are values beyond 1.5*IQR of each site of the box, and lines represent the range within 1.5*IQR."*
*We will refer to this explanation also 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. The labels of all used samples will be checked and labelling will be corrected to be in accordance with the labels that are used in PANGAEA to enable linking data.*

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 agree that the current plots are difficult to read in terms of interpretation of actual values and are thankful for the suggestion to change the labels. We will implement this in all plots where logarithmic scales are used to represent the data, as shown in the figure below as an example.*




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 understand that this is currently not clearly formulated in the manuscript. This number is the overall median (all calculation types), as given in Tbl. 4; the other two values are from the different calculation methods (calculation after Lončarić et al. (2006) vs. defining the BPZ as the end of the depth of the transition zone). We realised that the value for the range end is the mean in the current version of the manuscript, and not the median. The overall median value is 116 m. We will now give the two different values from the different calculation methods in brackets and explain their meanings.*

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 will also add 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 will implement in the text stating that is the lower end, so there is an increasing calcification intensity with increasing depth of 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 will include 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 agree and will include relevant values in brackets to make it easier to follow the arguments.*

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.

*Thank you for this suggestion which can add further information to our study. Such calculations can indeed be made for our data from the eastern Greenland Sea, eastern Norwegian Sea and both the eastern and western Fram Strait. We have now compared the abundances of planktonic foraminifera at those regions: Sampling with coarser mesh sizes (100 µm, 125 µm and 150 µm) results in abundances that are on average 27 % lower than abundances based on data sampled with a 63 µm mesh. Estimates of foraminifera abundances from data sampled with a mesh size of 125 µm and 150 µm are on average 69 % lower than estimates using samples taken with a 100 µm mesh. Only in the western Fram Strait, the abundances of N. pachyderma based on data taken with a mesh size larger than 63 µm appear to be higher (28 %) than based on samples taken with a 63 µm mesh. As abundances based on sampling with mesh sizes > 100 µm represent only 56 % of the abundances based on sampling with a 100 µm mesh in the same area, we assume that the higher value when sampling with a coarser mesh size is more likely caused by local influences on the foraminifera abundances, e.g. the climatic situation, and not on the method of sampling.*
*We will incorporate these values in the discussion in connection to the already cited values by Carstens et al. (1997), as an explanation on possible differences in estimated fluxes (new sentences marked in bold):*
***"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 samples 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 only 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 will change this.*

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

Line 572: contribution of planktonic foraminifera
*We will change this.*

Line 572/573: and a higher
*We will change 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. Whilst we prefer to remain with this coarse classification for the text, we suggest to include in the discussion section a text that explains that we are aware of the differences within the region, and that these can also have effects on the planktonic foraminifera standing stock and habitat, as shown in Pados & Spielhagen (2014), and formulate more clearly that the aim of this part of the discussion is to get a rough impression on the differences between regions in terms of foraminifera mass flux. We further checked the contribution of foraminifera fluxes to total $CaCO_3$ fluxes in more restricted regions, looking at the eastern Fram Strait (longitude > 0°) and the western part of the Greenland Sea (longitude < 0°) only. The results are not far apart, resulting in 10 % contribution in the eastern Fram Strait (in contrast to the before calculated 7 %) and 50 % in the Greenland Sea (in contrast to the before calculated 47 %), and we propose to use those values in the discussion to illustrate the potential effect of variations within the regions (changed parts in the paragraph marked in bold):*
*"A direct comparison of fluxes from samples from within the same region* ***with total $CaCO_3$ fluxes*** *in the region indicates a lower contribution in the* ***Eastern (> 0°)*** *Fram Strait (**10** %) and higher contribution in the* ***western part (< 0° E) of the*** *Greenland Sea (**50** %**; < 0° E**).* ***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 will change this.*

Line 581: than in other
*We will change this.*

Line 606: Arctic Ocean
*We will change this.*