**Reply to comment on essd-2022-324** (https://doi.org/10.5194/essd-2022-324-RC1) by the referee Lukas Jonkers.

The authors thank Lukas Jonkers for the time on reading and reviewing the manuscript. We sincerely acknowledge all the Reviewer suggestions that aimed to improve the manuscript readability and usability of the BENFEP database.

Our responses (plain text) to point-by-point comments are listed after the Reviewer comments (italic and bold font) followed by the main changes to the manuscript (italic font). Please, refer to the annotated manuscript and companion files.

***Section Introduction: The motivation to focus on the East Pacific remains somewhat unclear. The processes and Variability that the area is exposed to/characterised by (L31-39) do not seem to be unique to the Eastern Pacific. It would hence be good to be clearer about the scientific motivation to build a database for this specific region.***

We agree with the Reviewer´s comment. The Eastern Pacific is unique in that it holds the largest permanent oxygen minimum zone (OMZ) of the oceans. Climate change is currently leading to the expansion of the OMZ which make the Eastern Pacific Ocean especially vulnerable to diversity and habitat loss (as indicated in the references in the text). A synthesis effort to integrate benthic foraminifera geographic distributions would be beneficial for the stakeholders aiming to assess the vulnerability of this ocean area as well as the specialized scientific community aiming to construct modern analogues for the interpretation of the fossil record, in particular those related to paleoO2 reconstruction based on benthic foraminifera.

In the revised version of the manuscript, we have indicated the OMZ as a unique feature of the Eastern Pacific and clarified the scientific motivation to build a database for this region.

Changes to the text: Text rephrased (section 1, Introduction).
*In an historical context, the increase in ocean temperatures and the expansion of the already existing extensive oxygen minimum zone (found about 100 to 900 m water depth, Karstensen et al., 2000) are the major threats to shallow and deep-water benthic ecosystems from the Eastern Pacific (Sweetman et al., 2017; Breitburg et al., 2018; Yasuhara et al., 2019). These attributes make the Eastern Pacific an area of interest for assessing the past, present, and future of the marine ecosystem status and its response to expected environmental changes (e.g., Calderon-Aguilera et al., 2022).*

***Section Introduction: The authors also mention other syntheses. What is the scope of linking them? I understand that this is no trivial exercise, but I think a few lines somewhere (in the outlook) might be beneficial to the manuscript.***

We have added a few lines at the beginning of the penultimate paragraph of the introduction to link another synthesis. There, we highlighted the use of synthesis databases for developing proxies to interpret the fossil record. We also mention the use of those databases to assess the biodiversity change under ongoing climate change and to give the readers some background information about how extensive, both in number of samples and microfossil groups, are the existent databases of planktonic organisms, compared to the benthos.

Changes to the text: Text added (section 1, Introduction).
*The data synthesis of marine microfossils from surface sediments has been a valuable resource among the palaeoceanographic community. They are generally used for constructing modern analogues to interpret the fossil record and, more recently, to evaluate the biodiversity response to ongoing climate change (e.g., Jonkers et al., 2019; Yasuhara et al., 2020).*

***Section Taxonomy: …..Still the data is complex, perfectly illustrated by the fact that the authors recognise 1071 taxa, yet their dataset contains 1502 columns with abundance information. Because of this it would be good to provide a bit more detail on how taxa were mapped onto the WoRMS taxonomy. Was this done using the synonyms provided in WoRMS, using expert knowledge, something else?***

***Whilst looking at the synonym list and the taxon names in the data file I noticed the occurrence of many variants and subspecies, which are treated as separate categories/taxa. This of course risks inflating biodiversity and renders comparison of different studies challenging as subspecies and variants need not be consistently recognised. It would be good if their meaning and how they are treated would be described in the methods of the harmonisation.***

Thanks for your comments. As a requisite for data synthesis is the taxonomic harmonization of the given microfossil group, we decided to strictly follow the taxonomic criteria and synonymy included in WoRMS. We acknowledge that WoRMS is not the only taxonomic possible register for foraminiferal taxa but it is openly accessible to the public and it is continuously updated, which is an asset for a purportedly live database as BENFEP (now renamed BENFEP_v1).

We acknowledge the Reviewer's concern about the incomplete information about the harmonization procedure. Consequently, we have implemented the following changes that we hope to help to bring more transparency to the harmonization procedure:

- We have added a few sentences in section 2.4 to clarify the taxonomic harmonization procedure, as suggested by the Reviewer. There, we explain that we maintained the assignation given in WoRMS, even if the valid name resulted in a variety or subspecies. Additionally, we have increased the taxonomic information of each harmonized species in the Supplement section.

Changes to the text: Text added (new section 2.4, Taxonomic harmonization, previously section 2.5).
*In order to find the valid species name, we searched each author´s original species assignment in the WoRMS research engine. This procedure enables to identify whether the original species name is accepted (valid species) or if it is a synonymous of the valid species or taxa correspond to a variety or a subspecies. When the original species name was not currently in use, it was substituted by the valid species, subspecies, or variety name.*

The number of varieties (14), subspecies (4) in BENFEP_v1 is low compared to the number of total valid taxonomic assignations (1093: including identifications to species and below species level and the two taxa of uncertain status).

Changes to the Supplement: Please, refer to the new Supplement section.
The new Supplement File 1 incorporates more comprehensive taxonomic information (status, rank, order, family, class, AphiaID, etc) about the species harmonized in BENFEP_v1 than the previously submitted file.

Changes to BENFEP_v1 files:
The species numbers we presented in the submitted version have changed in the new reviewed version because we have received additional datasets (Mallon, 2011; Glock et al., 2020; Tetard et al., 2021) and we have included them. Besides, we have added the living counts of Smith (1964, 1973), Patterson et al. (2000), removed some samples with zero total abundance (wrongly included in the previous submission), and we have detected a few errors which are now amended.

***Section Taxonomy: …..In addition, could part of the reason for percentages summing to >100 be related to the reporting of species and subspecies without removing either the subspecies or the species prior to summing? This is a common issue in planktonic foraminifera datasets.***

Thanks for the remark. We reviewed that this is not the case for the species in BENFEP_v1. The percentages summing more and less than 100% were a source of continuous concern during the building up of the database. It is a common feature when data are sourced in percentages. It should be assumed that when the original sources reported the full (all species counted) dataset as percentages, the total sum should be 100%. However, in a few datasets that were digitized or manually digitized (in most of the cases these were reviewed by at least two people) the original datasets do not add to 100%. Therefore, we interpret that the original sources must have errors. In any case, we cannot discard some human errors from our side.

We think that sourcing the data in percentages is prone to errors and precludes some diversity calculations. For this reason, we make a plea to authors to provide raw data in counts (section 4 "Recommendations").

***Section Taxonomy: Some of the taxa appear to be fossil (e.g.https://www.marinespecies.org/aphia.php?p=taxdetails&id=927451), how are these treated? Are they real and therefore indicative of bioturbation/sampling of old material, or do they reflect taxonomic confusion?***

That is right. Indeed, BENFEP_v1 includes 109 taxa identified to species-level considered "fossil only" by WoRMS. The "fossil only" taxa are treated as the considered extant species. We justify our decision with the following arguments:

Firstly, we wanted to provide the users with an unabridged version of the original data. We followed the same criteria for any other non-numerical data originally provided by the authors (section 2.2).
Secondly, there are several reasons why a benthic foraminiferal species taken from a surface sediment sample might appear as fossil: a) It is an actual fossil specimen. In that case, the species might be reworked from nearby older sediments; b) it could represent a mistaken identification in the original source or c) the species is not a "fossil only" species but it has been inaccurately attributed as such by WoRMS. The latter is subject to change as new knowledge emerges; therefore, we consider important to retain the information as it appears in WoRMS. Eventually, this study might lead to revisiting the "fossil range" of some species.

In the new version of the manuscript, we have added the following information that could help the users to find whether a species is extinct or extant (according to WoRMS):

-In new section 2.4 we have now indicated that harmonized species include those considered "fossil only" according to WoRMS.

Changes to the text: Text added (new section 2.4, Taxonomic harmonization, previously section 2.5).
*Some taxa included in BENFEP_v1 are considered as "fossil only" by WoRMS. Nevertheless, we kept those in the database. There are several reasons to explain the occurrence of a species categorized as "fossil only" in a sample; it represents a true displaced fossil species from ancient sediments (reworking), a mistaken identification, and, an extant species inaccurately attributed as "fossil" by WoRMS. As it is not clear which of these circumstances applies in each case, we decided to maintain the species to prevent information losses in case of future re-evaluation of the "fossil range" by WoRMS.*

- In the new section 3.6.4, we have indicated the datasets that specifically mentioned species potentially reworked from old sediments and the relative contribution of those species to the assemblage.

Changes to the text: text added (section 3.6.4, The representativeness of the surface sediment assemblages, previously section 3.6.3).

*One of the purported applications of BENFEP_v1 is to provide a quantitative estimate of recent benthic foraminiferal assemblages that could be later used in palaeoenvironmental interpretations (e.g., Fig. 6). The database integrates quantitative data obtained from oceanic regions with different depositional environments, sedimentation rates, carbonate preservations and types of assemblages, collected over different sampling years and using an array of sampling devices that might result in diversion from recent conditions. For example, dead benthic foraminifera obtained from surface sediments might not be representative of the surface if the sampling device fails to recover the sediment-water interface or sedimentation rates are very low. The 36% of the surface sediment samples were retrieved using different types of coring devices (gravity, piston, dart and Phleger corer, calculations using "dev_1"), which are sampling techniques that can cause perturbation or miss-sampling of the surface sediment (Weaver and Schultheiss, 1990). Since the studies included in our database did not date the surface sediment (except for Palmer et al., 2020), we cannot discard that some samples correspond to pre-Holocene conditions. The most comprehensive compilation of sedimentation rates from core-top samples is from the equatorial Pacific and shows highly variable values, ranging from 0.8 to 14.2 cm/ka (Mekik and Anderson, 2018), meaning that surface sediment samples in this region correspond to recent conditions (assuming that no perturbation occurred during sampling). Reworking, downslope transport and carbonate preservation might be other factors influencing the composition of the assemblages obtained from the surface sediments.*

*The presence of "potentially fossil" species reworked from ancient outcrops (see "remark_2" and Supplement File 4) is included in the datasets of Bandy and Arnal (1957), Echols and Armentrout (1980), Ingle et al. (1980) and Zalesny (1959). Still, they represent less than 5% of the assemblage. The contribution of specimens displaced specimens from shallower locations is also low, as indicated by Bandy and Arnal (1957); Ingle et al. (1980); Harman (1964), Pettit et al. (2013), Uchimura et al. (2017) and Zalesny, 1959). Finally, Pettit et al. (2013) in the Gulf of California, and Boltovskoy and Totah (1987) and Resig (1981) off South America in samples below the carbonate compensation depth, are the only authors mentioning poor preservation of calcareous benthic foraminifera.*

- Because reworking has been mentioned by some authors, we collated the indications of the original authors in "remark_2" column. The "remark_2" of BENFEP_v1 guides the users to look for specific information in the new Supplement File 4.

Changes to the Supplement: Please, refer to the new Supplement section.
The new Supplement File 4 collates the original authors´ indications about potentially reworked species and comments about some potential misidentifications.

-Indications about each species "fossil range" as indicated by WoRMS in the new Supplement File 1

Changes to the Supplement: Please, refer to the new Supplement section.
The new Supplement File 1 includes a column dedicated to the "fossil range" according to WoRMS. We renamed the column as "occurrence" to indicate whether a taxon is extinct or extant.

***Section Taxonomy: The authors have chosen not to preserve the original name and have provided a synonym list in the supplement. The exact changes to each data set are therefore no longer traceable. How do the authors enable/envision future updates to the taxonomy or application of a different taxonomic framework where species lumped here are recognised as separate taxa? And would it be possible to include this information somehow to make the taxonomic framework more flexible and future proof?***

We acknowledged the Reviewer's concern about the incomplete information about the original author's taxonomic concepts procedure which might preclude reusability. In order to make clear to the potential users the taxonomic concepts provided by each author, and bring more transparency to the harmonization procedure, we have now included the following documents in the Supplement section:

Changes to the Supplement: Please, refer to new Supplement section.
-The new Supplement File 2 includes the original author's taxonomic concept for each species harmonized in BENFEP_v1 and included in new Supplement File 1 (please refer to comment above).

-The new Supplement File 3 includes specific remarks on the harmonization procedure.

***Section Taxonomy: One way to address many of these issues is to convert the data to long format. This would allow different taxonomies to exist in parallel (and hence preserve the original). Moreover, by having species, subspecies and varieties in separate columns (perhaps even genus to capture sp/spp) the hierarchy of the taxonomy can be preserved in a way that makes it easier to work with because it facilitates easy grouping at different taxonomic levels. I don't claim that this is the only solution, nor that it is absolutely essential to rework the entire database, but it would be good if the authors reflect more on their methodology and describe how they dealt with these issues or how they would recommend users of the database to deal with them.***

It is our goal to reach all type of potential users interested in benthic foraminifera, regardless of their preferred method to handle large datasets. For those who use more conventional software to read spreadsheet-type files, we keep the previously submitted format (BENFEP_v1_short). Those users will have the original authors' specific assignations in a separate file (Supplement File 2).

For those who prefer to access data using another type of software (e.g., geospatial software), we provide the database in long format (BENFEP_v1_long). The long-format version integrates all the information

provided in the short format plus taxonomic information for each species (File 1), "occurrence" (File 1) and each author´s species assignations (File 2), as suggested by the Reviewer.

Please, refer to section 2.5 of the reviewed version of the manuscript (which integrates former sections 2.4 and 2.6). There we explain the structure of the two formats of the database. Please, see the new Table C1 and Table C2, where we detail the column names and column codes for BENFEP_v1_short and BENFEP_v1_long.

***Section Limitations: There is very little discussion about what "living" actually means or how it is deduced. I think this could be useful for the non-specialist.***
***…. And finally, if living really means living, could the database be used to trace shifts in the assemblages over time given that the database covers a long period of sampling? Would this be a potential use of the data that could be mentioned in the introduction?***

We acknowledge the Reviewer's suggestion. Therefore, in the new version of the manuscript, we have briefly indicated what *living* and *dead* means (section 3.3). We expanded on the issue of using living or dead assemblages for modern analogue construction or to trace changes in the assemblages over time (new section 3.6.4).

Changes to the text: Text added and rephrased (section 3.3).

*The Rose Bengal staining (Walton, 1952) is the only method used by authors to distinguish dead (non-stained) from living (stained) foraminifera at the time of sampling. Living plus dead refers to an assemblage were living (stained) and dead (non-stained) are counted together in the same sample.*

Changes to the text: Text added and re-phrased (new section 3.6.4, previously section 3.6.3).

*BENFEP_v1 includes information of living, dead, and living plus dead assemblages whose suitability for building recent analogues is under discussion among the scientific community. The use of Rose Bengal as "vital" staining could be controversial because attached bacteria or algae, decaying protoplasm of dead individuals might stain, resembling the staining of the protoplasm of a "true" living individual (see review in Schönfeld, 2012). However, it is still the most widely used method to distinguish "living" (stained) from "dead" (non-stained) foraminifera and it is considered reliable if used cautiously. It might be argued that only living foraminifera should be used to consider baseline studies (Schönfeld, 2012). However, it might also be considered that living assemblages represent a "snapshot" of the foraminifera living at the specific time of sampling and do not hold the time-averaged representativeness of the dead assemblages (Murray, 2000). Regarding all these potential concerns, we have incorporated a rich collection of metadata to BENFEP_v1 that can be used by the final users to evaluate data quality and to tailor the final output to their specific criteria.*

***Section Limitations: Conversely, little attention to how old dead could be. Some more explanation and reference to discussion in the literature would be good in this regard. Paragraph 3.6.3 goes some way, but does not cite much literature.***
***The same paragraph also mentions that core top samples may not be representative of the most recent sediment if the sediment water interface is not captured. With the majority of the samples collected using gravity corers this could be a real issue, especially when dead and living specimens are not separated. Some discussion is probably warranted.***

We acknowledge the Reviewer's suggestion. Therefore, we have expanded on the information collated from the publications about age estimates of the surface sediment, foraminiferal preservation, and the presence of reworked specimens in new section 3.6.4 (previously section 3.6.3). Please refer to the reviewed document and to a reply to comment above.

Moreover, BENFEP_v1 contains metadata dedicated to the type of devices ("dev_1, dev_2"), sampling interval ("interval") and type of assemblages ("assemblage") which could aid users in filtering the data following their own criteria.

***Is the preservation potential of all species equal? I imagine that calcareous species are prone to dissolution and that agglutinated species are not always well preserved. Some discussion/guidance on this matter would be good.***

As the Reviewer indicates, a few authors of the datasets included in BENFEP_v1 mentioned poor preservation of calcareous forms: Pettit et al. (2013) in the Gulf of California and Resig (1981), Boltovskoy and Totah (1987) in samples located at water depths deeper than 3700-4100 m. Indeed, BENFEP_v1 incorporates datasets from the lower abyssal zone and below the carbonate compensation depth in the Eastern Pacific which are dominated by species with non-calcareous walls (Smith, 1973, Resig, 1981; Burmistrova et al., 2007; Enge et al., 2012).

We have now mentioned the authors indicating deficient carbonate preservation in new section 3.6.4 (previously section 3.6.3). Please, see reply to comment above

Regarding the preservation potential of foraminiferal species, benthic foraminifera are less susceptible to dissolution than planktic foraminifera and dissolution susceptibility varies among taxa. However, with 1093 valid taxonomic designations, the ranking of taxa solubility falls beyond the scope of this study.

***Section Limitations: Usage notes***
***The database is very heterogeneous. There are differences in the sampling instruments, sampling depths, picking method and in the level of detail in the taxonomy (some studies appear to report subspecies and variants, others don't) and the data is sometimes non numeric.***
***This is of course not the authors' fault, but it poses some important limitations on/challenges for the use of the data. I therefore miss some usage notes or recommendations on how to deal with this heterogeneity.***

***For instance:***
***How to handle non-numeric data;***
***How compare samples analysed using different size fractions;***
***What is the meaning of empty cells;***
***What to do when assemblages don't sum to 100%;***
***How to derive meaningful indices of biodiversity given the differences in the level of taxonomic detail; etc.***
***Some of these questions are perhaps trivial to the authors, but not all users of the data will be benthic foraminifera specialists and they might benefit from more detailed user instructions.***

We acknowledge the Reviewer's suggestions, and willing to help the final users, we have added the following changes to the main text and files. Please, find explanations to the Reviewer´s comments below:

***Non-numerical data:*** The handling of non-numerical data depends on the user's final purposes. For example, if an user wants to calculate sample species richness or species presence, they could replace the non-numerical data with 1, as the non-numerical data mostly indicates the presence of a species. In the case of using the data for percentage calculations, we would suggest removing them as they cannot be translated into a particular numerical value.

We have added a new section (3.6.3) providing general guidance regarding non-numerical data.

Changes to the text: Text added (new section 3.6.3. Non-numerical data).
*There are 18 datasets which include non-numerical data ("x", "<1") in their records (see "remark_1"). Those data might interfere in the calculation of the relative abundances and some diversity indexes (e.g., Shannon Weaver). However, they provide useful information on species presence and therefore they are potentially useful for biogeography and calculations of species richness. General suggestions on how to manage non-numerical data in R can be found in Supplement File 5.*

Changes to the Supplement: Please, refer to the new Supplement section.

-The new Supplement File 5 includes some suggestions on how to load the short format of the database in R and perform basic operations (e.g., build up a percent file), and manage non-numerical data

***Empty cells:*** this was indicated in the former Table B1: "An unfilled field, indicates no information available. It applies to any column". We have added the meaning of an empty cell in new Appendix C section (Table C1 and Table C2) and we have indicated the meaning of an empty cell in the new reviewed version of the manuscript (section 2.5): "An empty cell in any column indicates that there is no information available"

Changes to the Appendix: Please, refer to the new Appendices section.
Table C1 and Table C2 (previous Table B1) have been re-done to include explanatory notes on column names and column codes of BENFEP_v1_short (Table C1) and BENFEP_v1_long (Table C2).

***What to do when assemblages don't sum to 100%:***

We do not have a straightforward answer for that. Unfortunately, it is a common pattern in samples sourced from percentage. Our preferred choice is keeping the original data, rather than recalculating each individual percentage to 100%. This specially applies when data come from assemblages where authors do not provide the full scope of species or those which contain non-numerical data. As we make the user aware of this situation (section 3.6.2 and "remarks_1", "remarks_3" of the BENFEP_v1 files), we leave up to the final user the choice of excluding these samples from further analyses, recalculating the abundances, or keeping the original abundance data.

***How compare samples analysed using different size fractions;***
***How to derive meaningful indices of biodiversity given the differences in the level of taxonomic detail; etc.***

We consider that handling species and subspecies and size fraction is the users´ choice. They might choose adding a subspecies to the parent species and select samples according to a specific size range. The number of taxa below the species level ("varieties" and "subspecies") is low (18) compared to the total number of taxa to species level (1073 plus 2 taxa of uncertain status). See new Supplement File 1.

***Data section_ The format (xlsx) is proprietary and thus therefore not strictly adhering to the FAIR data standards. I recommend making the data available as some kind of text file that is universally readable and more likely to remain so.***

Being in completely agreement with the Reviewer, and in order to support the reusability of digital assets, we now provide BENFEP_v1_short and BENFEP_v1_long as text files.

***Data section_ The species names in the headers appear inconsistently formatted and not to map one-to one to the names in the synonym list in the supplement (use of underscores and brackets). I realise that this is a pain, but it would enhance machine readability if the formatting were made consistent.***

We understand and agree that consistent formatting is a must. Consequently, we have checked that the headings of BENFEP_v1_short do match the headings of the species list in the Supplement (File 1, File 2).

***Data section_Another, very small remark, is the use of special characters in the column headers, which complicates (machine) reading and the authors could consider replacing them with plain text.***

We understand and agree that the use of special characters in the column headers complicates machine reading. Accordingly, we have made the following changes in BENFEP_v1 to facilitate usability. The changes are explained below:

Changes to BENFEP_v1 files:
-simplification of column headers: taxonomic designations are simplified to contain only the genus and species name (or varieties, subspecies) in plain text. Thus, we have removed information of authority, year

or any special character (e.g., parenthesis, &, etc). The users could find this information in Supplement File 1 or in BENFEP_v1_long in columns "valid_authority".

-removal of capital letters from column names and restrict column length to 10 characters to facilitate use in GIS software.

-use of spaces: albeit some exceptions (taxa names), we try to avoid using spaces between words in cells and column names. In case a space is required, we added an underscore.

***Minor comments***
***L10: consider adding "benthic" at the start of the sentence.***

Done as suggested

***L29: I don't understand what "figures" means here?***

We have re-phrased the text in that line to:

*Several areas of Eastern Pacific Ocean are at severe risk of species loss (Finnegan et al., 2015; Yasuhara et al., 2020, UNESCO, 2022) and consequently, some of them have been categorized as marine protected areas (Enright et al., 2021).*

***L58: reword, something like "rendering interpretation of the fossil record more meaningful."***

Done as suggested

***L171-172: I assume that the database can be handled in any software, especially once converted to text format. Why focus on these two examples only instead of highlighting that it can be analysed using any software on any platform?***

Thank you for the suggestion.
Because we provide BENFEP_v1 in a text file (which is a more universal and license-free format) we followed the Reviewer's suggestion and added the text accordingly.

Changes to the text: Text rephrased (new section 2.5 Structure of the database, previously section 2.6)
*The database in its two versions is presented in text format and can be managed with virtually any software.*

***Fig. 2 legend: the colour scale does not, I presume, show the log10(number of species). I think it shows the number of species as normal numbers, but with a logarithmic colour scaling. I would just remove the (log10).***

The reviewer is right. We have modified the figure in the reviewed version of the manuscript.

Changes to the Figures:
Figure 2 has been re-made to amend the legend and to update the changes made in the files (see previous explanation to Changes to BENFEP_v1 files).

***Fig. 4: perhaps this is a figure where I wonder what the exact purpose is. Do we learn much about the data and why is that useful?***

Changes to the Figures:
We have removed the former Figure 4 from the reviewed version of the manuscript.

***L223: these size ranges seem inconsistent with the figure below, those shown in the figure are wider. Why? And are the ">" signs needed?***

We are sorry that the text and the figure were not clear enough. We now provide an alternative version of the figure and rephrased the text accordingly. We believe this new plot is clearer than the former one.

Changes to the text: Text added and rephrased (section 3.3).
*Most of benthic foraminiferal assemblages were analyzed in the smallest size fraction commonly used in benthic foraminiferal studies. For example, 65.4% of the samples were analyzed using 42, 61, 62, 63 and 74 μm as the lower end of size fraction (e.g. assemblages where studied in the >42 μm, >61μm, >62 μm size fraction, etc.). The 6.1% is using 88 and 105μm and 17.7% the 125, 149, 150, 200, 212 and 500 μm as the lower end of size fraction.*

Changes to the Figures: new Figure 4 (previously Figure 5).
In the reviewed version of the manuscript, we have re-done Figure 4. We have calculated the relative contribution of each size fraction (>61, >62, etc) independently and represented it as a portion of the pie chart in the new Figure 4.

***Fig 7D: are there really samples with 0 species? And how exactly was this figure made, see my comments above on user instructions.***

This figure was built up by calculating the number of taxa per sample. In a few sampling points there were no foraminifera (0 abundance and 0 number of species). It was inconsistent to include those samples in the previously submitted version. Therefore, we have now removed those samples from the new files of the BENFEP_v1 database.

Changes to the Figures:
In the new reviewed version, we have re-plotted Figure 6D (previosuly Fig. 7D) with the information of the new BENFEP_v1 files.

***Fig. 7E: I am not sure in which direction the abundances are summed in this figure (rows, columns or both?). If the authors want to show the depth distribution of selected species then I would expect that each row should sum to 100%, but that seems not the case.***

The figure was built up by calculating the mean relative abundance of each of the selected species for each bathymetric range. We have now indicated that in the figure caption *(….and selected species mean relative abundance with water depth (E) in BENFEP_v1*)

Changes to the Figures:
In the new reviewed version, we have re-done Figure 6E adjusting it to the bathymetric ranges of Morkhoven et al., (1986). The bathymetric ranges we used in the submitted version were incorrect as they were designed for the pelagic environment.

***L287: "A few samples…" I counted 423 and 98 >105%, that is not a few. Just provide the numbers.***

Thanks for the annotation. This sentence was re-written to "well above 100%" and we indicated the entry it refers specifically to (Butcher, 1951). As we mentioned earlier, not summing 100% is a common feature of the data sourced in percentages.

***L331-332: something went wrong with the sentence "It is recommended It would be…"***

Thanks for the annotation. The sentence has been amended.

***L333-334: the authors should be bolder here, providing taxonomic information is crucial for the reusability of (any) taxonomic data. So it is not only of value to the "specialised foraminifera community" and unclear taxonomic information compromises the data and prevents their reusability. Using data only once (in one study) really is a waste of time!***

We have followed the Reviewer's recommendation and we have changed the text accordingly.

Changes to the text: Text rephrased (section 4 Recommendations, previously section 5).
*Taxonomic information and supporting images are crucial elements for reliable taxonomic harmonization and data reusability.*

***L340: I recommend some kind of versioning scheme so updates can be more easily recognised.***

The reviewer is right; we did not consider versioning properly.

We renamed the database file to BENFEP_v1 to differentiate it from future versions.
Section 5 contains now information about how we foresee versioning after BENFEP_v1.

Changes to the text: Text added and rephrased (new section 5. Data availability and future plans, previously section 6).
*The BENFEP_v1 database can be accessed from https://doi.org/10.1594/PANGAEA.947086 (Diz et al., 2022b). This database is conceived as a springboard to store future quantitative data of benthic foraminifera in the East Pacific and make them available to the scientific community. It will be open for any new quantitative data entry and thus, it welcomes any new data published or provided by any contributor. The database will be updated by the authors once a considerable number of new entries need to be incorporated or changes are required to update taxonomic categories to an existing version. New versions of BENFEP will be submitted and curated in PANGAEA. Collaborations with individual researchers and institutions are welcomed specially regarding potential expansion to other ocean basins.*