We thank the reviewer for her valuable comments and suggestions regarding our manuscript entitled "*Ecological response to collapse of the biological pump following the mass extinction at the Cretaceous-Paleogene boundary*". Below follows a point-by-point response to the comments by the reviewer. Comments by the reviewer are **in bold**, our reply is in normal font.

---

**[Research approach: it should be kept independent from the main models (e.g. Living Ocean model, see comment below) concerning the K/Pg marine biological crisis. The already known models must not be used to interpret the data, differently the reasoning gets circular preventing any new knowledge from emerging.]**

**--**

We disagree with the reviewer here. Naturally, the benthic foraminiferal and dinocyst data were obtained and analyzed 'independent from the main models'. However, when interpreting the obtained results, it is, in our opinion, crucial to view the results in the context of known models.

It should be noted that, although in the past decades there has been a considerable debate on the occurrence of a collapse of the biological pump following the K-Pg boundary mass extinction (e.g. Alegret and Thomas, 2009), almost all recent studies are in agreement that post-K/Pg export productivity was reduced *to some extent*. Current discussions mostly involve discussions on the *severity* of the reduction of export productivity, or on the geographical differences (i.e. the 'heterogeneity' of the oceans). Even studies that argue for a more 'heterogenic' response of the global marine ecosystems, such as Esmerey-Senlet et al (2015) argue that, although there was spatial heterogeneity in the wake of the K/Pg mass extinction, "*the interbasinal $\delta^{13}C$ gradient was reduced after the mass extinction, suggesting a reduction in global export productivity*".

Moreover, while records from for example the open Pacific Ocean could, potentially, be explained as showing no reduction of the biological pump (e.g. Alegret and Thomas, 2009; Alegret et al., 2012), there is no discussion on the occurrence of a collapse of a biological pump in the Tethys Ocean, Atlantic Ocean, Southern Ocean and Indian Ocean (e.g. following Thomas, 1990; Olsson et al., 1996; Hull et al., 2011; Alegret et al., 2012). We therefore do not see it as the target of our study to 'test' the living ocean model. Instead, given that it has been convincingly shown that this model is valid for the Tethys (e.g. Hull et al., 2011; Alegret et al., 2012, etc.), the main focus region of this study, the objective of this study was to investigate the potential ecological consequences of this early Danian 'living ocean' condition.

---

1. **[Title: The model arguing for a global collapse of the biological pump following the mass extinction is controversial, and still not univocally accepted (see**

**Thomas, 2007, Birch et al., 2016). I suggest to the authors to remove it from the title.]**
**--**

Most recent studies are in agreement that, *at most sites*, export productivity was reduced *to some extent*. Also the reference cited by the reviewer (Birch et al., 2016 "Partial collapse of the marine carbon pump after the Cretaceous-Paleogene boundary"), states "*Our results show that changes in ocean circulation and foraminiferal vital effects contribute to but cannot explain all of the observed collapse in surface to deep-ocean foraminiferal d13C gradient. We conclude that the biological pump was weakened as a consequence of marine extinctions, but less severely and for a shorter duration (maximum of 1.77 m.y.) than has previously been suggested*."

Indeed, there have been discussions on the geographical heterogeneity of the oceanic response (Hull et al., 2011; Alegret et al., 2012). However, even studies that argue for a more 'heterogenic' response of the global marine ecosystems, such as Esmerey-Senlet et al (2015) argue that, although there was spatial heterogeneity in the wake of the K/Pg mass extinction, "*the interbasinal $\delta^{13}C$ gradient was reduced after the mass extinction, suggesting a reduction in global export productivity*" and that there is convincing evidence for a collapse of a biological pump in the Tethys Ocean, Atlantic Ocean, Southern Ocean and Indian Ocean (e.g. Hull et al., 2011) Therefore, using benthic foraminiferal and dinocyst records from these regions (not the Pacific) can be used to assess the ecological response to the collapse of the biological pump in these regions. Therefore, we argue that it is valid to use this term in our title.

---

2. **[Introduction: Pag. 4-L9-13: this paragraph states the approach of this paper which in my opinion is conceptually wrong. You don't do carry out a new research to place it "in the context" of what it is already known or thought to be known, but to bring in new knowledge, improve, edit or discard what's already known.]**
   **--**

We agree with the reviewer that 'placing in the context' is not a goal on itself. Naturally, this was not the goal of this study. Therefore, we will rewrite the section accordingly.

---

3. **[Methods: The authors studied the size fraction larger than 125 µm for the benthic foraminiferal analysis. This can lead to miss important ecological information as disas- ter taxa and stress tolerant opportunistic taxa which bloom during environmental stress are often smaller (e.g., Boscolo Galazzo et al., 2013; Giusberti et al., 2016). To me the use of the >63 µm size fraction would have been more appropriate for this study. See for instance Thomas (1990), Alegret et al. (2003), Alegret and Thomas (2007; 2009). The study of**

**the smaller size fraction might for instance reveal peaks of small opportunistic infaunals, challenging the current environmental interpretation. Ideally the counts should be improved counting the whole >63 um size fraction. I understand that at this stage this would imply the re-study of the whole sample set. However, the authors should at least re-count same samples using the whole >63 um size fraction in order to check that important ecological information/patterns are not missed in the critical stratigraphic intervals with the use of the larger size fraction. These additional data should be included as a figure in the paper. ]**

**--**

We fully agree with the reviewer that studying the size fraction >63 um is generally preferable, especially in deep-sea environments with many small taxa. However, the preservation of the studied material from Okçular is rather poor, resulting from strong recrystallization during deep burial of the sequence. Already in the >125 um size fraction, accurate determination of foraminifera even to the genus level was often difficult. This resulted in a relatively high 'Indet.'group. A study on the smaller size fraction would provide better insight into the abundance of small taxa, yet this would go hand in hand with an even higher indeterminable fraction, further compromising an accurate portrayal of the benthic foraminiferal assemblage. It should be noted that the comparative data of the similar but much better preserved material of El Kef is also based on the >125 μm fraction.

Furthermore, we are not fully convinced that the choice of size fraction necessarily influences the overall patterns significantly. This is exemplified by the study performed on the PETM by Ernst et al. (2006). In this study, the authors compared the results of the separately picked size fractions (>63 and >125 μm). Although there is indeed additional information on smaller taxa in the 63-125 μm fraction, the overall patterns prove to be robust. This suggests that it is useful to compare assemblage changes (not absolute numbers!) found in our study with those in other studies, even when the size fractions used in the studies are different.

---

**[To estimate benthic foraminiferal accumulation rates (BFAR) in on-land sections can be somewhat difficult as sample dry bulk density values are difficult to measure. In this work average density values derived from literature are used For this reason, I advise caution with the use of these BFAR data to reconstruct export productivity changes, and I recommend BFAR is not used as a key parameter to interpret benthic foraminiferal faunal changes.]**

**--**

We agree with the reviewer that estimating the BFAR in on-land sections can be difficult as sample dry bulk density values are difficult to measure. However, in our study the influence of variations in dry bulk density between different stratigraphic intervals on the estimated BFAR will be minor, as the densities of mudstones, siltstones and claystones generally all fall within a range of 2-2.5 g/cm$^3$. In the Okçular section, the

BFAR shows a decrease of <u>almost 2 orders of magnitude</u>. This is such a large decrease, that the influence of any minor variations in dry bulk density will be negligible. Nevertheless, to exclude the possible influence of the (small) differences in dry bulk density, we will reflect the uncertainty in BFAR, resulting from the uncertainty in dry bulk densities, in our Figure 5.

---

**[Besides, they calculated BFAR using the number of benthic foraminifera/gr sediment for the >63 um size fraction while their assemblage counts have been done in the >125 um size fraction. This must be changed for consistency as faunal patterns in these two size fractions can be quite different.]**

**--**

The reviewer has a valid point in this respect. We agree that the BFAR should have been estimated using the number of benthic foraminifera/gram sediment for the >125 um size fraction for consistency. We have now done so for our dataset and will plot the >125 um size fraction BFAR in Figure 5 and will provide both the >63 um size fraction BFAR and the >125 um size fraction BFAR in our Data Set S2 (supplementary materials).

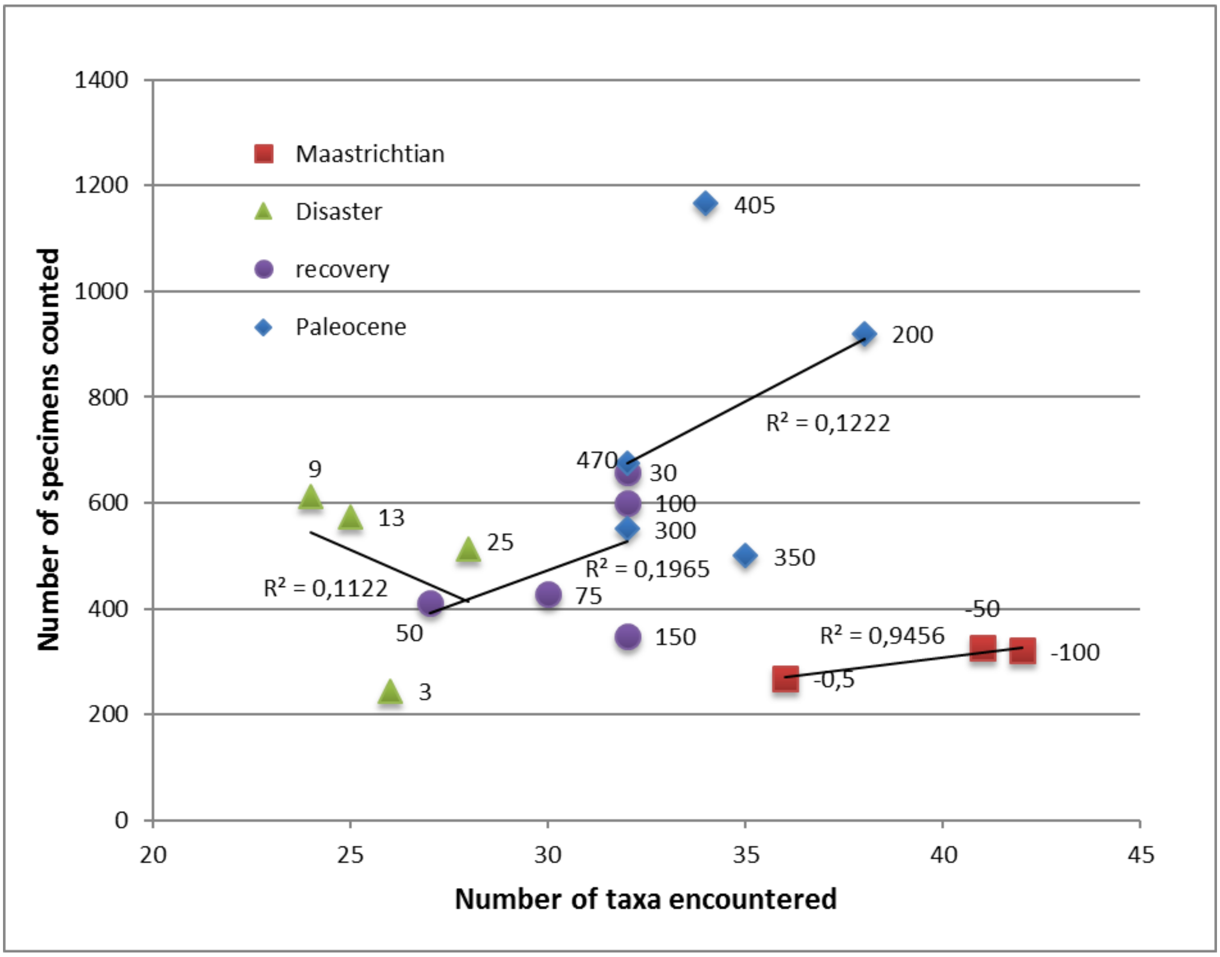
---

**[Benthic foraminiferal counts have been made by counting 300 specimens for each samples. This is a standard counting threshold widely used in benthic foraminiferal quantitative studies. However in this specific case I encourage the use of species-specimens plots to establish the most suitable number of specimens to count (see Thomas, 1990). The use of species-specimens plots allows to ensure that species diversity is well represented. In an outer-neritic upper bathyal site species diversity might be higher than in the deep sea settings, where the standard average of 300 specimens is usually employed. Since the relevance of diversity changes among benthic foraminifera for this study I would perform a species-specimen plot for each of the 4 intervals recognized in order to assure species diversity is well represented.]**

**--**

The reviewer argues that the number of specimens per sample may be too low to capture the full diversity, since in an outer neritic-upper bathyal site diversity may be higher than in the deep-sea. First of all, we note that a selection of >250 specimens per sample is a standard procedure, both in deep-sea and continental margin studies (e.g. Murray, 2006). Some authors recommend a higher number, e.g. counts of 300-400 specimens (Lowe and Walker, 1997). In this study, only in two cases a slightly lower sum than 300 was obtained (245 and 269 specimens, respectively). In all other cases, the sum was well over 300, and often over 500 (Figure 4 in the manuscript).

In other studies on deep-sea and outer neritic benthic foraminifera across the K-Pg boundary, for instance the study from the Bidart section (Alegret et al., 2004), the Walvis Ridge section (Alegret et al., 2007), the Agost section (Alegret et al., 2003), and the El Kef section (Speijer and van der Zwaan, 1996), the counted number of specimens is mostly 200-300. Compared to these studies, the Fisher alpha (which calculated from number of taxa and total specimen count per sample) and Shannon H diversity are lower in our study, also in the cases where the counted number of specimens is higher than 600. As Fisher-alpha should in theory not be sensitive to number of specimens counted, especially when a relatively high sum of >250 is reached (see also the charts in Fisher et al., 1943), this is a strong indication that diversity in our study is indeed systematically lower than in the other sites mentioned, instead of higher. Finally, if the number of species encountered in each sample is plotted against the number of specimens counted per sample (figure below, to be added to the SI), we see no trend towards higher numbers of species with increasing sample size. We thus argue that we do not have any indication that the number of specimens in our study may be too low to capture the full diversity.



*Number of specimens counted vs number of taxa encountered, for the 4 different stratigraphic intervals. Both within each stratigraphic interval, as well as between the different intervals, there is no clear relationship between number of specimens counted and number of taxa encountered.*

---

**[4. Results: please mention in the text (1) the number of samples along with their stratigraphic position for each of the recognized intervals (benthic foraminifera); (2) thickness and approximate duration of each interval.]**

--

We are aware of the complications that arise when benthic foraminiferal data, dinocyst data or other assemblage data are plotted in graphs that are completely black, the 'filled silhouettes'. In such graphs, it is impossible to determine the sample depths from the graph and to determine the number of samples used to create the graph. For this reason, we decided to use colored silhouettes with horizontal black lines at each sample depth. This way, it becomes clear immediately how many samples were studied in each interval, and at which depths these samples were taken. Adding numerical data on the number of samples studied and their stratigraphic position in the figures is therefore, in our opinion, unnecessary. However, if the reader would be interested in exact sample depths and numbers of samples used, they are referred to the SI where all information is outlined in detail. In addition, we will indicate the recognized intervals in our Data Set, so that it becomes clear which samples belong to which interval. It should be noted that both the stratigraphic depth and biostratigraphic tie-points are indicated in the figures. This allows the reader to address both the thickness and approximate durations of each interval.

---

**[5. Discussion: -Paragraph 5.1: As highlighted by the review paper of Culver (2003) and by Thomas (2007) there is no agreement regarding the ecological meaning of the generally low-diversity benthic foraminiferal assemblages occurring just above the K/Pg boundary. Even though similar changes between the % of epifaunal and infaunal species can be recognized between different records, such % changes can have different environmental meanings in different environmental settings. So I suggest caution in drawing Tethys-wide environmental scenarios based on changes in the proportion of epifaunal-infaunal species. ]**

**--**

We agree with the reviewer that changes between the % of epibenthic and endobenthic taxa can have different environmental meanings in different environmental settings. Especially when one compares shelf sites, such as in the Tethys, with deep marine, open ocean sites, such as in the central pacific, the ecological meaning of, for example, high abundances of endobenthic taxa can be different. However, particularly for this reason, we mainly compare sites in the Tethys ocean, characterized by similar palaeodepths (upper bathyal/outer neritic), similar palaeolatitudes (25-35°) and similar sedimentologies (mixed carbonate-siliciclastic sedimentation), which, therefore, represent similar environmental settings. We therefore argue that Tethys-margin-wide environmental scenarios *can* be drawn based on the recorded signals in these records.

---

**[Paragraph 5.2.1: In this paragraph the authors seem to use the Living Ocean model (D'Hondt and Zachos, 1998) to explain their data. In my opinion they**

**should first provide a sound interpretation of their dataset and then, argue whether their dataset fits (or not) with the main models used to explain the K/Pg δ13C shift. ]**

**--**

The first paragraph of section 5.2.1 (lines 11-17) describes our dataset. To address the comment of the reviewer, we will also include a couple of sentences to provide an interpretation of our dataset, independent of the Living Ocean model. The argument whether these data fit the main model, will then follow this sentence, in the second paragraph of section 5.2.1

---

**[Further, which new contributions brings their own data to a further development/understanding of these models? In my opinion this is an aspect which is currently not sufficiently addressed in the paper.]**

--

We do not fully understand this comment of the reviewer. In section 5.2.2, we describe what new insights our study brings, showing that, as a result of the reduced biological pump efficiency, more nutrients became available for the earliest Paleocene phytoplankton community. Our study is the first to suggest that at Tethyan neritic to upper bathyal sites this increased nutrient availability is reflected by the higher abundance of hexaperidinioids (Lines 28-34).

---

**[Paragraphs 5.2.2&5.2.3: As for what concern benthic foraminifera, the work cited in these paragraphs used different size fractions for their studies (either >63 um or >125 um) so direct comparisons among datasets are so far not possible. ]**

**--**

See point 3.

Note that in our study, the comparison is drawn between the Okçular and El Kef datasets, both of which are based on the >125 µm fraction. In addition, the study by Ernst et al. (2006) has convincingly shown that, although there are minor differences between >63 µm fraction and >125 µm fraction datasets, the overall biotic patterns prove to be robust. This suggests, in contrast to the reviewer's view, that comparisons between overall assemblage changes between different studies are possible, even when different size fractions are used.

---

**[6. Conclusion: The conclusion paragraph should be focused on summarizing the findings of the paper. Personally I think it should be rewritten highlighting the paper's data and their meaning. This is first of all a paper which presents new**

**ecological data from a new section spanning the K/Pg boundary, it is not a review paper.]**

**--**

We differ in opinion with the reviewer on the structure of the conclusion. Over the last decades, numerous studies have been published on K-Pg boundary benthic foraminiferal and dinocyst records, from an array of sites around the world. The most novel, exciting aspect of our study is not that we are adding an extra benthic foraminiferal and dinocyst study to this collection. Rather, it is the combination of our new ecological data with existing records in the same geographical region that provides us valuable new insights in the paleoecological and paleoceanographic consequences of the early Danian 'Living Ocean' condition in this region.

---

**[7. Figures 2-6: please add the number of samples studied and their stratigraphic position. ]**

**--**

We used colored silhouettes with horizontal black lines at each sample depth. This way, it becomes clear immediately how many samples were studied in each interval, and at which depths these samples were taken. Adding numerical data on the number of samples studied and their stratigraphic position in the figures is therefore, in our opinion, unnecessary. However, if the reader would be interested in exact sample depths and numbers of samples used, they are referred to the SI where all information is outlined in detail.

---

**[Figure 6: please add the duration of each interval like in the previous figures.]**

OK. We will include the stratigraphic tie-points in Figure 6, similar to the previous figures.

---

**[Minor remarks: Abstract: Pag. 1-L26: beginning of the line, please insert "in" after comma.]**

OK

---

**[Text: Pag. 2-L13: "toward" repeated twice. ]**

Corrected

---

**[Pag. 4-L11: "records" repeated twice. ]**

Corrected

---

**[Pag. 5-L22: please delete “refractory”.]**

OK