Response to Reviewer 2 Comments (RC2):

Irvali and co-authors present a very interesting study with new ideas for establishing improved chronologies of modern sediments. Such efforts are much needed in paleoceanography because better chronologies for the upper, most recent sediments will allow the comparison of proxy data with instrumental time series and can therefore contribute to improving proxy calibrations and reconstructions.

The new data presented in this study include measurements of stable isotopes on foraminifera to detect the anthropogenic Suess effect and the occurrence of SCPs as additional age controls for the last few centuries.

Although the research idea and the data produced are of high quality, I have some major concerns about the chosen methods, the reporting of the data and structuring of the manuscript.

We appreciate the reviewer's thoughtful feedback and suggestions. We address the specific comments below.

In the introduction, the authors list various available techniques for dating recent marine sediments, and state that all of them have their own limitations and uncertainties. With so much data available for the studied sediment core, the ideal approach would thus be to combine all information into an integrated optimized age model for the last few centuries. However, the authors choose to create a core chronology based only on stable carbon isotopes measured on one species of foraminifera and their correlation to a model (based on the Suess effect), complemented by radiocarbon dating. The radiocarbon dating uses a reservoir age which is obtained from the abovementioned Suess correlation, so this is not independent either. The available Pb-210 and Cs-137 information is discarded and only included in the discussion, and the new SCP data is not actually used to build the age model but only to confirm the findings based on the isotopes.

Ultimately we agree, this manuscript is a case study useful for moving us in that direction, and specifically exploring the utility and consistency of using 13C-Suess. Once the utility and limitations are well defined, then a more integrative approach is certainly merited, although much of the information available is qualitative to semi-quantitative in nature.

On lines 108-109, the authors argue that the Cs-137 concentrations below 4 cm core depth were too low to detect and therefore these data were not used. One could argue that this is a result in itself; the anthropogenic Cs-137 isotope only occurs above 4 cm. I would consider this valuable information for the age model.

Thanks for this suggestion. We will fix and clarify this in the revised MS. The results from the Pb-210 and Cs-137 measurements from Mjell et al., 2016 *(their Supplementary Figure S1)* are shown below. In the case for core GS06-144-09MC, we think that Cs-137 is difficult to use unambiguously. Cs-137 is not only present above 4 cm, but also episodically below this depth (see below). As a result, it is not clear what this represents in chronological terms and so we consider it not useful to include in the age model.

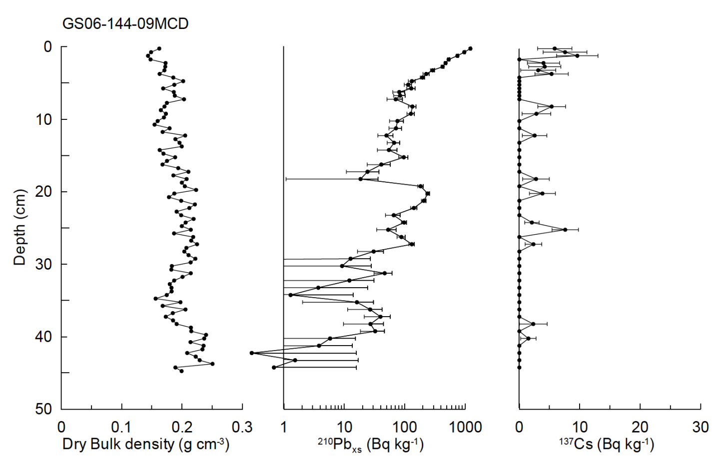

Figure 1. Down-core variability of dry bulk density, unsupported ^{210}Pb content, and ^{137}Cs content in core GS06-144 09MC-D (Mjell et al., 2016).

The same is true for the radiocarbon age of the top of the core (Table 1, date KIA34242). The authors correctly state that this age indicates the presence of bomb carbon, and therefore younger than ~1957. This is another constraint that can be used to include in the age model. I don't understand why it is "not included in the age model". It is not possible to use it as a normal radiocarbon date, but an age constraint in the form of a maximum age should be possible to incorporate in the core chronology.

We thank the reviewer for bringing this up. This will be clarified in the revised MS. As the reviewer suggests, the core top radiocarbon date provides a valuable information and confirms that the core top is modern (0 cm >1957) (Line 235). However, here, we have instead used 2006 (the year the core was retrieved) to provide a maximum age and set a limit when building our age model in BACON, as the core top age (0 cm) cannot be younger than 2006.

The interpretation of the Suess effect in the foraminiferal isotope measurements is steered towards the final conclusions of the authors and does discuss other options. Only *bulloides*, and the combined stack are compared to the model output (Suppl. Table 1), but not the other individual species? Why not?

We prefer to use *G. bulloides* and the δ^{13}C stack (to cross-check the results we obtain from *G. bulloides*), instead of using all individual species (*N. incompta* and *G. inflata*). The main reason behind this is to look at the combined signal rather than habitat specific shifts and to avoid making further assumptions regarding foraminiferal habitat depths (e.g., particularly for *N. incompta*, which has a more variable depth habitat), and believe that using an average of 0-200 m would cover the habitat depths of all the planktonic foraminifers used in this work.

The curve fits are compared to various depth intervals from 12-0 to 5-0 cm (Lines 203-204), but why stop there? What about the curve fit coefficients at 4-0 cm, or even up to 1-0 cm?? The best fit is found for 7.5- 0 cm with r=0.73. Is this significant compared to the other intervals? For 0-5 cm, the r value is 0.69. Is that difference significant? And what would it be for 0-4 cm etc?

We appreciate this suggestion. We will extend the correlation analysis to the topmost part of the core in the revised MS and explore the possibility of including another test to evaluate the statistical differences between the curve fits for the various depth intervals.

After this simple statement: "7.5 cm must be 1800 AD", this is taken as fact, without any further discussion or even consideration of uncertainty. In SI Fig. 3 or in Fig. 4, one could imagine a very good fit between the red and blue curves if the red was more compressed and shifted to the right.

To make a more objective match of these two curves, and, instead of visual / wiggle matching, we have used polynomial curve fits and correlation analysis. Then, 7.5 cm is selected as it is based on the highest correlation coefficient between the Suess effect-*G. bulloides* comparison. Hence this statement is based on correlation analysis results presented in Supplementary Table 1.

The core top age is then determined as 1977, based on the findings above and a very simplified comparison between the upper parts of the isotope model data, and the *bulloides* measurements (Figure 3). This figure, instead, to me highlights the main difference between the data. One change is extremely abrupt, showing the major changes all in the top samples, while the other is gradual, with a decrease over 200 years. This difference is not sufficiently discussed. Also, this figure illustrates how the biggest change in *bulloides* is all in the top 1 cm of the core, and the moving average curve is here left out.

We will include a more detailed discussion of this. The moving average curve will also be included in the revised figure.

If the core top would actually be 1977, where is the rest of the sediment? Multi-cores are known to preserve the sediment-water interface, so what happened to the last 34 years?

The multicore GS06-144-09MC was retrieved in 2006, and if the core top is 1977, we would be 'missing' the last 29 years. Taking our average sedimentation rate for the 0-7.5 cm interval (~43 cm/kyr or 0.043 cm per year) at face value, we could then expect ~**1.2 cm** of sediment to be deposited in 29 years. However, the fact that we don't have a "modern" core top (or 1.2 cm more sediment) suggests changes in sedimentation rates, bottom flow speeds or bioturbation to be likely candidates.

From here onwards, this new age model is simply taken as the truth and no uncertainty is reported whatsoever. The term "known-age" is described between quotation marks, but that is not the same as discussing or reporting uncertainties. At some point (Line 345) a ±3-year uncertainty is reported, but it is not clear where this value comes from?

The authors claim to be able to deduce better estimates of the local ΔR values with lower uncertainties. They briefly acknowledge possible bioturbation, but don't discuss that the new ΔR values are based on the assumption that 0 cm is 1977 and 7.5 cm is 1800. The uncertainties derived from this should be included in the new reservoir age estimates.

We appreciate the reviewer' suggestions. We will state and clarify our assumptions in the revised MS.

Figure 7 shows a combination of some of the data, compared to the old Pb-210 age model. The very short discussion that follows just repeats how the new age model was made but fails to explain why the Pb-210 doesn't match. What could be reason for the mismatch? And would it not be possible to find a solution that satisfies all data? Perhaps the raw Pb-210 data could be reevaluated, and not just the resulting age-depth model.

We will include a more detailed discussion to explain this in the revised MS.

The manuscript often lacks a clear distinction between introduction, methods, results, interpretation or discussion. An example is paragraph 3.5 on SCP analyses. It combines most of the above in a single page.

Paragraph 3.5 will be rephrased to correct this in the revised MS.

In summary, the presented research idea is very promising, but in the current state, the methods and discussion do not make a convincing case that the new chronology is more reliable than the old one.

Specific comments:

- Lines 44 – 86. Several studies on recent marine sediments have used the increase of mercury concentrations as an anthropogenic marker for the last century. This could be added to your list. Example of a study from north of Iceland: https://doi.org/10.1371/journal.pone.0239373

We thank the reviewer for this suggestion. This will be included in the revised MS.

- Lines 116-117: *bulloides* was picked from the 250-300 µm size fraction, while *inflata* was picked from the 250-350 µm size fraction. Is this a typo, or did you include a 300 µm AND a 350 µm sieve?

This is correct. A narrower size range for *G. bulloides* is preferred as it is known to show variable stable isotope values based on differences in shell sizes. Hence, *G. bulloides* was only picked from the 250-300 µm fraction, while *G inflata* was picked from both 250-300 µm and 300-350 µm fractions.

- Line 155: "We set the starting point in time to 1800,...". How does the record look before that? Does it still look similar to the measured isotopes?

G. bulloides $\delta^{13}C$ record vs the atmospheric $\delta^{13}C$ from Rubino et al. (2013) is plotted below, spanning the last 1000 years . However, we avoid a direct comparison due to the sparsity of the data/sampling resolution before 1800.

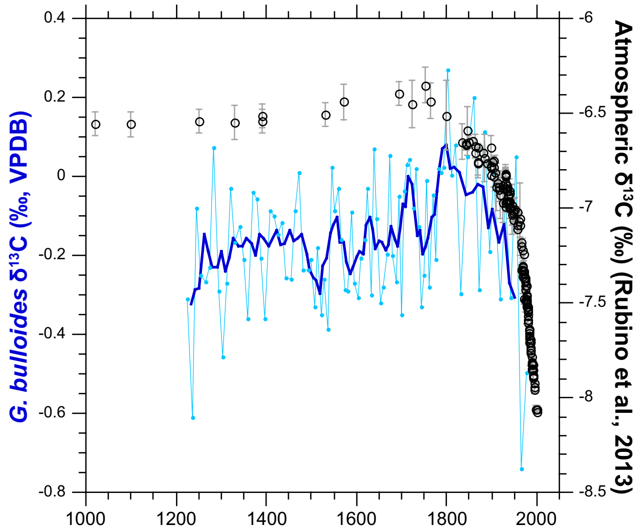

Figure 2. *G. bulloides* $\delta^{13}C$ record from Site GS06-144-09 MC-D *(blue, with 5-point mean)* plotted together with the atmospheric $\delta^{13}C$ record of Rubino et al., 2013.

- Lines 191-193 does not give an objective description of the isotope results. Natural variability is described as "over the 10-44 cm core interval" which already includes the interpretation that thereafter, the changes are of anthropogenic origin. Why not explore the option that "natural variability goes to 4 cm? Or 1 cm depth? Why stop at 10?

We agree. The natural variability exists throughout the core (0-44 cm), and it would be impossible to distinguish this or state that there is no natural variability. In L191-193 we only suggest that there is a <u>larger</u> natural variability over the 10-44 cm interval of the core; while the anthropogenic influence increases and dominates the most recent interval/the core top. This has been discussed in more details in Lines 401-410. However, as also suggested by Reviewer 3, we will highlight more strongly the potential sensitivity of our approach to high natural variability in the revised MS and explore the possibility of including a sensitivity study on various cases.

- Lines 210-211: "suggesting 7.5 cm must be 1800 AD". This is too strong a statement. What about the uncertainty of this method and any critical discussion?

We will highlight the uncertainty of this approach in the revised MS.

- Lines 230-231: "gives us a rough estimate of which curve is most similar to our target curve (i.e., d13CSE_0-200), and overall agrees with our initial finding". Again, this is simply not critical enough. Would it also have agreed if you had other findings? Probably.
- Line 287: "the ΔR in the region is highly variable". What is meant here are the ΔR values based on the online database of calib.org, which are just a few observations. It is not the same as the actual ΔR values, so this should be clarified.

We agree. We will fix wording to clarify this in the revised MS.

- Line 323: Bioturbation is not limited to the top 10 cm of the core. Every depth level was once the top of the sediment, so bioturbation affects the entire core. The mixing or resulting smoothing of data is then over a ~10-cm window.

We will fix wording to clarify this in the revised MS.