Dear Editor and Reviewers,

We thank you for your positive feedback and the critical evaluation of our manuscript. We have taken each comment and suggestion into account. Below we provide our responses (in blue) to each individual comment/suggestion and describe how they were incorporated. We also note that we added a new figure 1 showing a schematic of the Kalman Smoother reconstruction process. We are confident that the revised version is significantly improved and look forward to your feedback.

Sincerely,

Sönke Dangendorf

(on behalf of all co-authors)

Reviewer 2:

This paper provides a new probabilistic reconstruction of historical global and regional sea-level changes by combining two methods (RSOI and Kalman smoothing) and using recently published results to constrain the reconstruction. Overall, the paper is clearly written, the methods used are well explained (although some improvements could be made I think - see below) and the results bring new information on historical sea level, which makes this paper and the related dataset definitely of interest for future research and operational studies. I only have a list of minor comments, suggested corrections and questions that I have listed below.

We thank the reviewer for their positive and critical evaluation of the manuscript.

- Please include more carriage return in the introduction to make it easier to read.

  The template format might have given the impression that parts of the manuscript are one long paragraph, while in reality they are not. In the final print lines 47-97 should appear as 3-4 paragraphs

- P2l48: 3.4 mm/yr over which period ? please clarify.

  Added (since 1993)

- P4l122: “control input parameter” or “control input parameters” ?

  Thank you, corrected

- P4l123: please revise the equations (1) and (2). The “k” subscript on the right hand side should be replaced by “k-1” in (1) and the “O with stroke” in (2) should be replaced by phi (state transition matrix) I presume.

  Thank you for spotting this, corrected!

- P4l128: “the latter […]” please reformulate. I don’t really understand what is the first adjustement compared to Hay et al.

  We adjusted the sentence, which now reads as follows:

"*The reduction of the vector to only one globally uniform GMSL estimate is the first modification to Hay et al. (2015, 2017), where the vector also contained estimates of 21 melt rates from mountain glaciers and ice sheets.*"

- P4l130-132: were all individual GRD estimates used from Frederikse et al. (2020) or the total contribution (i.e. 1 GRD estimate for all components)? Please clarify.

  We added "*total*", but we also note that the selection of the prior estimates is discussed in the data section. There we also added a sentence "*Here, the sum of the individual contributions is considered.*"

- P6l171: replace 4,0000 by 4,000

  Thank you, done!

- Although this not done either in Hay et al. (2017), I'd suggest to add a flowchart Figure that illustrate the various steps of section 2.1 to guide the reader through the methodology and also underline in this flowchart, where the present work differs from Hay et al. (2017) (and Dangendorf et al., 2019) to better capture the novelty.

  Thank you, we added a schematic as Figure 1 that illustrates the main equation in combination with the data that goes into them. We do not see how differences to Hay et al. (2015) and Dangendorf et al. (2019) can be easily incorporated. Those, however, have been clearly outlined in the text.

- What is the time step of the procedure? annual? Please clarify.

  This information is provided in the introduction as well as the next sentence (first sentence of data section), so we decided to keep this as is.

- P7l213: " […], which we interpret as sterodynamic" (?)

  Not sure what the issue here was but we adjusted the sentence as follows:

  "*The resulting sterodynamic sea-level fields are then used as input parameters for the Kalman Smoother.*"

- Figure 2: please recall in the caption the time period over which correlation is calculated for clarity.

  Thank you, done

- Figure 4: I've the impression that the mismatch between the present reconstruction and Frederikse et al. (2020) is well pronounced before ~1990, while from 1990-onwards, the comparison seems good. Could that be the effect that the RSOI-based reconstructed sterodynamic component is strongly constrained by the availability of altimetry data? Please comment.

  The two curves match very well over most of the record with the exception of the behavior in the 1930s and 1940s (the rest is largely spurious local variability from the reconstruction process which is larger in Frederikse et al. (2020) due to the averaging technique that they used). Both techniques are not directly comparable. Frederikse et al (2020) is based on a virtual station technique (i.e., averaging) and does not use any satellite altimetry, so we cannot see how the use of satellite altimetry should influence their (dis-)similarity. We also note that it is only covariance information that comes from

satellite altimetry, while tide gauges are still the main source of information in the reconstruction.

We had a global sea level reconstruction intercomparison project a few years ago (https://www.issibern.ch/teams/unifysealevel/) in which we tested the different reconstruction techniques in terms of their performance in climate model fields in which the model truth is a priori known. The conclusions from that intercomparison (not yet published) were that the handling of VLM information was the main source of uncertainty between different techniques (which is in line with results presented in Dangendorf et al., 2017) with advantages towards field reconstructions (virtual station techniques can be strongly biased by individual gauges with datum shifts or other local processes, that would be counted as outliers in the field reconstructions).

- P14l340-342:"This limitation is evident […] smaller amplitude signal in the tide gauge record of San Fransisco". This statement is unclear to me. I rather see a larger amplitude signal in the TG record (?). Also, figuring where the 1997/1998 El Nino signals pops-up is not straightforward.

  We adjusted the text as follows:

  "*This limitation is evident by comparing satellite altimetry observations and the reconstruction during individual peaks such as the 1997/1998 El Nino where the reconstruction underestimates the amplitude compared to the tide gauge record at San Francisco (**Figure 6a**)*"

- P15l367-369: GNSS datasets used should be described in section 2.2 (Data). I supposed the ULR7A solution from Gravelle et al. (2023) was used? If yes, please clarify this too (as four solutions are available on SONEL actually).

  Yes, we used URL7A and added that information. We only included data for the reconstruction itself in the data section and introduced validation data (as GNSS and the other steric products) in the main text. This ensures that the text remains brief and avoids any unnecessary double mentioning.

- P15l381: -0.14 or -0.13 ? Please make the main text and Figure 6c consistent.

  Done, thank you

- P17l415: Maybe it would be relevant to cite the corresponding Figure of Frederikse et al. (2020). Fig1c right ?

  The record is shown in Figure 7b (now Figure 8), so it is unclear why the Figure in Frederikse et al. (2020) should be cited.

- Figure 7: colors in the plot and the legend for Frederikse (2020) budget don't match. Please correct this.

  Corrected, thank you!

- P18l445: please refers to Figure 9 by quoting Figure 9 already there.

  Done, thank you!

- Figure 9: the Sterodynamic component rates suggests ~30yrs oscillation. Is that possibly related to climate variability or is it related to the reconstruction processing? Such an oscillation is not obvious in e.g. Frederikse et al (2020) (see their Figure 1c)

This is hard to tell given the huge differences between individual observational products and the large uncertainties in that component within our reconstruction. For the moment we prefer not to emphasize any of that given the reconstruction uncertainties.

- P21l473: "2.2 +/- 0.11" while Table 3 gives 2.2 +/- 0.12. Please make the main text and Table information fully consistent.

  Done, thank you

- P22l477-479: here also main text and Table 3 rates estimates are not always consistent. Please adjust

  Thank you, we forgot to update the numbers in the text but have adjusted them now.

- P23l495-496: Figure 11a colorscale does not allow appraising such larges ranges as it ranges from -3 to 3 mm/yr. Maybe modify the caption or adjust the main text to clarify this.

  Thank you. The difficulty is to emphasize the main features in the different contributions. As a solution, we added information to the colorbars in Figure 11 and 12 (now 12 and 13) that the values can exceed or fall below the color range.

- P23l497-498: "are therefore by the GIA signal". A word is probably missing (?).

  Thank you, we adjusted the sentence as follows:

  "*Negative trends are primarily located in centers of postglacial uplift where former ice sheets were located (Scandinavia and North America) and are therefore caused by the GIA signal and to lesser extent due to GRD…*"

- P23l515-516: "The local maxima […] reflect gyre dynamics". unclear. Reformulate ?

  We adjusted the sentence as follows:

  "*The local maxima around ~40 degrees North and South are due to the sterodynamic contribution and reflect gyre dynamics.*"

- P24l519 : Would it be worth and reasonably feasible to test a different SLP dataset ? E.g. HadSLP? or in light of what is done for GIA, using various ensemble members of the 20CR reanalysis dataset?

  The choice of using only one SLP dataset was motivated by the fact that an additional dataset would double the ensemble size, which makes the computation too heavy. We also note that the overall contribution by inverse barometer effects is very small compared to the other components. Therefore, we decided to stick to one dataset in this version but aim at testing multiple datasets in future releases. Please also note the response to a very similar comment made by the other reviewer.

  We have adjusted the methods section, which now reads as follows:

  "*IBE variations are derived from sea-level pressure data from atmospheric reanalysis models, specifically the 20th century reanalysis v3 covering the period 1900 to 2015 (Slivinski et al., 2019) and the NCEP-NCAR reanalysis 1 (Kalnay et al., 2018) from 2016 to 2021. Both reanalyses were interpolated from different global grids onto the 74,742 grid points for the regional sea-level fields considered in the Kalman Smoother. They are combined after adjusting them to the same mean at each grid point over the five overlapping years between 2011 and 2015. The choice of using the 20th century*

reanalysis v3 is based on its availability over long periods. Other datasets (e.g., HADSLP) are not considered, as the IBE has only a very minor contribution compared to other processes and every additional dataset would double the total ensemble size considered here."

- P28l608-610: this information should be made very clear in the introductory part too I believe.

  We added that information also to the end of the introduction, which now reads as follows:

  "*Second, the approach enables the estimation of the individual contributors and sea-level change as constrained by tide gauges. As such we derive a novel global and regional mean sea level reconstruction at annual resolution covering the period from 1900 to 2021 extending the reconstructions from Hay et al. (2015) and Dangendorf et al. (2019) by 11 and 6 more years, respectively*"

- P29l638 typo 'aimpacted'

  corrected