Reply on Review 2 (Kappelsberger et al., 2023)

Thank you very much for this supportive feedback on our manuscript. Your suggestions, such as including a flowchart visualising the entire workflow, are very helpful in improving the presentation of our manuscript. Below we respond to each comment and describe how we plan to revise the manuscript to make it shorter and easier to follow and read. Reviewer comments are marked in italics.

1 Major concerns

The manuscript is very long and for a long time, I was unable to see where we were going and how all the notation, techniques, differences etc. were to be used. I realize that the authors want to be systematic by introducing all the methodology in Section 3 before using it in Section 4 and discussing it in Section 5. But this means that, until somewhere in Sections 4 and 5, I had been loaded with a large amount of notation, techniques etc. without really knowing why I needed to know this.

1.1 *Try to shorten the manuscript. I know you want to be thorough and systematic, but is there really no way of making it shorter? It is a very long read.*

- We fully agree that the paper will benefit greatly from being shortened. The following parts of the manuscript will be significantly shortened or even deleted:
  - L55-63 (introduction)
  - L66-70 (introduction)
  - L72-79 (introduction)
  - L91-101 (introduction)
  - L195-231, Fig. 2 (data)
  - L320-326 (methods)
  - L414-419, Fig. 6 (results)
- In addition, we plan to change our methods slightly for ease of presentation and explanation.
  1. Data section: Here, we will describe that we remove the offset, linear, quadratic and seasonal signals from the altimetric elevation changes, h^{A}, and from the modelled firn thickness changes, f^{M}, to obtain the altimetric variations, hv^{A}, and the modelled firn thickness variations, fv^{M}. Thus, hv^{A} and fv^{M} will represent our "input data sets" for our methods below.
  2. Methods section: Based on this, we will use hv^{A} as the input to our regression approach. This will simplify Eq. (1) to

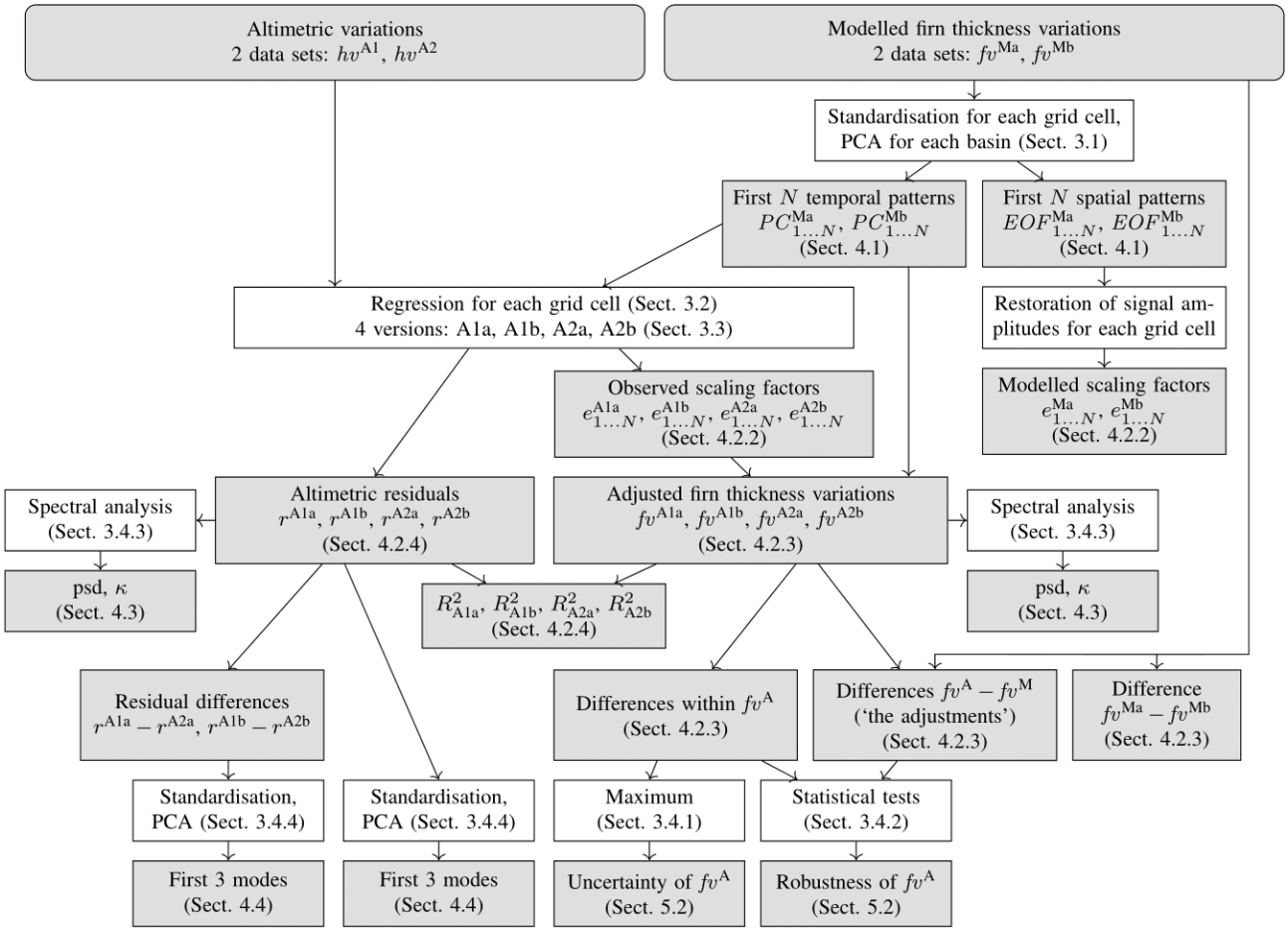
$$hv^{\mathrm{A}}(t) = \sum_{n=1}^{N} e_n^{\mathrm{A}} \, PC_n^{\mathrm{M}}(t) + r^{\mathrm{A}}(t).$$

  Thus, Eq. 2 and its accompanying text will no longer be needed at this stage as we will remove the offset, linear, quadratic and seasonal signals before. Likewise, Eq. 4 and its text are also no longer needed at this stage. We might change the order of the subsections 3.1 and 3.2 due to the planned changes.

  We note that by splitting the regression approach (Eq. 1) into two separate steps of parameter estimation (one for the offset, linear, quadratic and seasonal part, and a second one for the interannual patterns), we will imply a slightly different treatment of correlations between those parameters than in our original manuscript. However, the effect on the results will be small. Our conclusions will not be affected. With these changes the methods section will be shorter and we believe, much easier to read and understand.

1.2 *Somewhere in the beginning make an overview of what the problem is, and the pathway to solve it. Try to include a cartoon or flowchart showing the data coming in, all the intermediate products, the residuals, the analyses done on the residuals, and the use you make of these things. Include also the notation in each box of the chart and the relevant section numbers - not just where they are derived in Section 3 (as in Table B1) but also where they are calculated in Section 4 and used in Section 5. That would provide a road map for the reader making the journey through it all easier to navigate in.*

- This is a very good idea, thank you. At the beginning of the methods section we will include a flowchart showing the whole workflow. The flow chart might have the following structure:



Here, the (intermediate) results, their notation and where they are presented for the first time (Section 4/5) are shown in grey boxes. The main methodological steps (Section 3) to derive these (intermediate) results are shown in white boxes. Please note that the (sub)section numbers in the flowchart do not correspond to those of the submitted manuscript, as we plan to slightly change, swap or combine some (sub)sections.

2 Minor issues and typos

L53: SOME RCMs specialize…

- We will rephrase: "When the main goal of RCMs is to realistically simulate the ice sheet weather, as is the case here, they are forced by atmospheric reanalysis products (van Wessem et al., 2018; Agosta et al., 2019)."

L178: ERA5: What is the resolution of ERA5? 31 km, right? And this is "downscaled" to 27 km, right? Please argue why? Do you have reason to believe the 27 km RACMO data is better than the 31 km ERA5 data as input to the FDM?

- Yes, the resolution of ERA5 is 31 km. ERA5 is "dynamically downscaled" to 27 km by RACMO2.3p2 (van Wessem et al., 2018). Lenaerts et al. (2019) describe this method as follow: "(…) dynamical downscaling, that is, to use a high-resolution regional (atmospheric) climate model (...) forced at its boundaries with (...) reanalysis data. This ensures a sufficiently high resolution over the ice sheet as well as an explicit, physically realistic calculation of all SMB components."

  Thus, RACMO2.3p3 has two advantages over ERA5. First, the spatial resolution is improved. An improvement from 31 km to 27 km does not sound like much, but for each grid cell it already means a 25 % reduction in area. By improving the spatial resolution, processes such as katabatic winds or orographic precipitation can be better simulated in complex, mountainous regions (Hansen et al., 2021). This leads directly to the second benefit: the improvement of simulating physical processes for the special case of polar regions. RACMO2.3p2 is coupled to a multilayer snow model (there is an interaction between the atmosphere and snow surface), snow albedo is considered, a drifting snow scheme is included (van Wessem et al., 2018).

L181: MERRA2 is downscaled from what resolution to 12.5 km? And how?

- MERRA-2 has a resolution of 0.625° longitude x 0.5° latitude, giving a resolution of 24 km x 56 km at a latitude of ±70° (Medley et al., 2022). Medley et al. (2022) used a downscaled version of MERRA-2, which they called "replay" MERRA-2 run. This replay run was produced within the NASA Downscaling Project (Tian et al., 2017).

<u>In the revised manuscript, we will remove subsection 2.3, to which the following three comments refer:</u>

- *L204: The use of parentheses and the two short sentences "Rates…" and "The three…" is quite clumsy here. Please rephrase.*
- *L206: "evaluated by comparing to"*
- *L214: Try to think of a better title for this sub-section*

L219 (and many other places, e.g., 401): The use of parenthesis after a full stop is not the usual way of doing it. Either a parenthesis refers to and is part of the sentence that is full-stopped, or maybe it should not be a parenthesis at all.

- We will go through all these sentences, remove the parentheses and rephrase the sentences if necessary.

L262: The sentence starting with "Observations" is very difficult to read. Perhaps start with the fact that you find higher noise levels from the older sections and use this to motivate why you introduce a weighting. Also, say which variable (r^A?) represents this noise you talk of.

L265-269: This is very difficult to understand. Please see if you can rewrite it more clearly.

- We will rephrase the whole paragraph and also consider the comment on L321. The following proposal for reformulation already refers to our proposal for the modified methodology, where hv^A, rather than h^A, applies as observations of the regression approach: "The stochastic model of our regression is represented by a different weighting of observations from two time periods. As results from the older altimetry missions generally have a higher noise level (Schröder et al., 2019a; Nilsson et al., 2022), hv^A after 2003 are weighted by 1, while hv^A before 2003 are given a different (usually lower) weight. The weight before 2003 is defined (individually for every grid point) by the ratio of the noise variance of hv^A before and after 2003. We assess the noise by the high-pass filtered version of hv^A separately for both periods (cf. Groh et al., 2019). The high-pass filtering consists of removing a low-pass filtered version of hv^A, where the low-pass filter is a Gaussian filter with a 6σ = 12 months filter width."

Eqn 1, 2, and 4: These equations all include a, b, c, d1, …, but they are different (and subject to different regressions) in the three equations, right? Either change the notation or write this out very clearly.

- Eq. 1 and 4 are two different regressions. Eq. 2 is not subjected to a different regression, only some of the signals estimated by Eq. 1 are removed. However, to simplify the methodological presentation, we plan to apply Eq. 2 as a separate regression before fitting the dominant temporal patterns in modelled firn thickness variations to altimetry (see the explanations above under 1.1).

L321: "deterministic": What do you mean by deterministic? And later on (L669+673), you talk of "stochastic". Exactly what is stochastic? I cannot see any noise terms added anywhere in your methodology.

- In general, the mathematical model of a least-squares adjustment consists of (1) the deterministic or functional model and (2) the stochastic model given by the variance-covariance matrix. Here, the deterministic model is given by Eq. 1. The stochastic model is not given by equations, but it is described in L262-268. The term "weighted regression" means that our stochastic model is only a diagonal matrix, so only variances and no covariances are considered. The term 'stochastic model' should be clearer with the reformulation of L262-268.

L371-375: Was very difficult to read. I think I understood it when the results were shown later on, but when reading it here I did not get it.

- To make that part easier to understand we plan to include concrete matrix sizes: "For each PCA, we set up one aggregated 'super data matrix' in which we arrange the time series of residuals/residual differences for grid cells and for the different versions into a single set of time series. Specifically, our data sets comprises m = 90638 points in space (entire area under investigation) and p = 108 points in time (2003–2017). Thus, for the first and second PCA, the super data matrix has the size of 4m x p and 2m x p, respectively. The PCA is conducted to identify the dominant temporal patterns (PCs), which are shared by all versions, together with their space-dependent and version-dependent amplitudes, i.e. their spatial patterns (EOFs). Each identified mode thus consists of one joint PC (1 x p) and four/two EOFs (4m x 1/2m x 1) in the case of the first/second PCA."

L380: You say that the fv^Ma are standardized prior to PCA, but in Fig 4 you say that the EOFs have units of m. How can the EOFs have units if the input is standardized and thereby non-dimensional?

- Thank you very much for pointing this out. Indeed, since fv^M are standardized prior to PCA, the EOFs should not have a unit. We will remove the unit of meter in Fig. 4.

Fig 7: Why are d-f not identical (or at least similar) to the EOFs in Fig 4? Are you not projecting the model signal on to the PCs that came out of a PCA on exactly that signal? Should that not be a way of recovering the EOFs, i.e., by projecting the signal onto the PCs? Or does it have to do with standardized vs non-standardized signals?

- Exactly, it has to do with the standardisation. In the revised manuscript, we plan to explain it in a different way: "To regain interpretable magnitudes of the EOFs, the EOFs are multiplied by the corresponding std of the time series of fv^M for each grid cell (which was previously used for standardisation). After this restoration of the signal amplitudes, we no longer speak of EOFs but of modelled scaling factors, e^M".

  Also, we will note in Fig. 6: "(d-f) is the same as Fig. 4a-c but with restored signal amplitudes for each grid cell." In other words, Fig. 4a-c is the standardised version of Fig. 6d-f. The patterns of Fig. 4a-c and Fig. 6d-f are similar, but as the standardisation is done for each grid cell, this may not be obvious.

L455: What is R_s? I cannot remember having this introduced before.

- True. It should be R^2.

L478: "underlying time series IS displayed"

- Here, we wanted to refer to both time series, fv^A and r^A. To clarify that, we will write: "The underlying time series of fv^A and r^A are (…)".

L635: The sentence starting with "Thus, the …" is difficult to understand.

- We will delete this sentence as it repeats the previous one.

L712: The sentence starting with "We deliberately" does not read well. Particularly the word "deliberately" seems odd. Please try to rephrase the sentence.

- We will rephrase the first and second sentence of the paragraph: "We developed a new approach that combines satellite altimetry and firn modelling results to resolve Antarctic firn thickness variations at a high temporal (monthly) and spatial (grid scale of 10km) resolution."

L723: outperforms

- We will rephrase: "The adjusted firn thickness variations, fv^A, outperform …"

L735-736: The sentence starting with "Across basin 8" makes it sound as if you only did this spatial analysis over basin 8, but didn't you do it over all basins?

- Yes, the analysis was done for every grid cell but particularly in drainage basin 8, we found differences in the spatial patterns between fv^A and fv^M. We will delete this sentence.

L736: Suggest to combine the sentence starting with "The large" with the one before.

- As we delete the previous sentence, we will rephrase: "The large uncertainty in basin 8 is likely due to the presence of megadune fields".

L746: Do you not rather subtract the modeled firn thickness variations from the altimetric variations? That is what eqn A1 says, but your text says the opposite.

- Thank you, yes, it should be: "we simply subtract the modelled firn thickness variations from the altimetric variations".

L791: Where can the IMAU-FDM data be found?

- The IMAU-FDM is on GitHub and Zenodo: https://github.com/brils001/IMAU-FDM or https://zenodo.org/records/5172513. For Antarctica, the data can be also provided by the authors. We will include this information in the revised manuscript.

All the following comments will be revised as suggested:

- *L294: "scale it SUCH that"*
- *L481: "HAS stronger"*
- *L482: "IS closer to"*
- *L601: includes*
- *L691: "in THE snow"*
- *L725+731: The sentences "However, one caveat should be noted." are a bit odd and short. Suggest you combine them somehow with the sentences coming after.*
- *L733: resolveS*
- *L734: "evaluated AT grid cell level"*
- *L735: Perhaps underline that the basin 5 and 8 numbers are also calculated at grid cell level as in the previous sentence.*
- *L736: "are due" should perhaps be "are likely due" or some other modifier to weaken the claim.*
- *L745: "TO the original"*

Additional references:

- Hansen, N., Langen, P., Boberg, F., Forsberg, R., Simonsen, S., Thejll, P., Vandecrux, B., and Mottram, R.: Downscaled surface mass balance in Antarctica: impacts of subsurface processes and large-scale atmospheric circulation, The Cryosphere, 2021, 4315–4333, https://doi.org/10.5194/tc-15-4315-2021, 2021.
- Tian, B., Lee, H., Waliser, D. E., Ferraro, R., Kim, J., Case, J., Iguchi, T., Kemp, E., Wu, D., Putman, W., and Wang, W.: Development of a Model Performance Metric and Its Application to Assess Summer Precipitation over the U.S. Great Plains in Downscaled Climate Simulations, J. Hydrometeorol., 18, 2781–2799, https://doi.org/10.1175/JHM-D-17-0045.1, 2017.