Author response to the review of Anonymous Reviewer #2

Johannes Landmann and co-authors

June 2021

Dear Anonymous Referee #2,
Dear Editor,
We would like to thank Anonymous Referee #2 for the valuable review of our manuscript. We appreciate the constructive comments and the positive feedback regarding the overall context of the study. We address the comments point-by-point answers to the individual remarks and questions.
Best regards,
Johannes Landmann and co-authors

– Major comments and questions –

RC: *What made you choose a particle filter as DA method as opposed to more traditional methods (e.g. variational methods and ensemble Kalman filter). It would complement the work if you discussed why a PF suits this problem.*
AR: The Particle Filter (PF) is a generic data assimilation method that can handle all kinds of model distributions, also non-linear ones. We have chosen the PF since we know that the distributions we deal with are not always Gaussian. To not extend the already long introduction any further, we suggest to add these explanations at l. 250, where we introduce the Particle Filter: "Especially when temperatures are around the melting point, the system becomes non-linear since melt occurs above but not below this point. As a consequence, the distributions we deal with are not necessarily Gaussian. The facts that (a) the temperature chosen to parametrize the melting point is not the same for all four models, (b) the individual model prior distributions are combined to obtain the ensemble prediction, and (c) there can also be accumulation contributing to the overall mass balance, add further complexity. We do not use other data assimilation approaches, such as variational methods or Ensemble Kalman filtering, because variational methods encounter difficulties when dealing with non-Gaussian priors (van Leeuwen et al., 2019), whilst the Ensemble Kalman Filter in its original form is not designed for multi-model applications as we use in our case. Overall, particle filtering is a very flexible, generalizable, and readily implementable data assimilation method."

RC: *Line 105. If I understood well, the observations are of a cumulative quantity. In this case, do observation errors need to consider time auto-correlations?*
AR: This is a justified question. No, the observations are not cumulative, in the sense that the mass balance at a given time is not inferred by summing the individual, sub-daily readings up to that time. Rather, the mass balance of a given point in time is given by one single reading at that time. In this sense, the individual measurements are independent, and only affected by the precision by which we can read a stake at a given moment. To clarify this, we suggest adding the following text at L.107: "We expect the observation errors to be uncorrelated in time, since every reading is independent from the previous one."

RC: *Line 255.I was a bit confused on where the uncertainties of the input variables are represented. Are they represented in the model error beta, in the observation error epsilon (as mentioned in line 258), or both?*
AR: The observational error from the camera readings is represented in the observation error ϵ, as stated in l. 257. The model input errors are considered to be contained in the model error β_t, as stated in l. 258. We suggest to better clarify this by replacing "[...] (β_t) can also represent uncertainties in model input variables" with "[...] (β_t) should include the uncertainties about model input variables" in l. 258.

RC: *The use of the PF in a multi-model ensemble context is quite interesting, especially since each model has different parameters one is trying to estimate. Is there previous work in this regard? Could you*

provide some references?
AR: Currently, we are not aware of other studies that have applied particle filtering in a multi-model ensemble related to glacier mass balance. However, we have added some references to applications in other contexts:

- Kreucher, C., Hero, A., & Kastella, K. (2004, March). Multiple model particle filtering for multitarget tracking. In Proceedings of the Twelfth Annual Workshop on Adaptive Sensor Array Processing.
- Ristic, B., Arulampalam, S., & Gordon, N. (2004). Beyond the Kalman filter: Particle filters for tracking applications (Vol. 685). Boston: Artech house.
- A. Saucan, T. Chonavel, C. Sintes and J. Le Caillec, "Interacting multiple model particle filters for side scan bathymetry," 2013 MTS/IEEE OCEANS - Bergen, 2013, pp. 1-5, doi: 10.1109/OCEANS-Bergen.2013.6608125.
- Wang, R., Work, D. B., & Sowers, R. (2016). Multiple model particle filter for traffic estimation and incident detection. IEEE Transactions on Intelligent Transportation Systems, 17(12), 3461-3470.

The revised text will read: "We are not aware of mass balance studies that have applied a multi-model ensemble based on a particle filter with the resampling methods we propose, although multi-model particle filters have been used for other applications (e.g. Kreucher et al., 2004, Ristic et al., 2004, Saucan et al., 2013, Wang et al., 2016)."

RC: *When discussing the particle filter, you introduce the concept of 'minimum contribution' for some particles. This is taken into account when weighting, as explained in appendix 2. There is a comment saying that the original weights are preserved 'in average'. Could you elaborate more on this statement?*
AR: Only preserving the weights on average is common for resampling procedures: when a particle performs poorly, it obtains a weight of zero and disappears. If a particle $x_{t,k}$ has a weight $w_{t,k}$, then it is chosen $N \cdot w_{t,k}$ times in the resampling. This particle then has the weight 1/N times the "number how often it has been resampled", so on average $w_{t,k}$. The same is true when resampling within a model: a particle $x_{t,k}$ with model index j is resampled on average $N_{t,j} \cdot w_{t,k}/\pi_{t,j}$ times. After resampling it has the weight $\tilde{w}_{t,k}$ times "number how often it was resampled", so on average $w_{t,k}$. This statement is made in Equations B2 and B3, which are found in Appendix B.

RC: *Equation 21. How are μ_0 and Σ_0 chosen?*
AR: We choose μ_0 and Σ_0 from the parameter statistics obtained from the calibration procedure in section 3.2. We suggest to add a phrase stating where μ_0 and Σ_0 originate from. Suggested revised text: "$\vec{\mu}_0$ and $\vec{\Sigma}_0$ are the prior mean and the prior covariance of $\vec{\theta}$ at the starting time t_0, which we determine from the calibration procedure described in section 3.2,[...] "

– Minor comments and questions –

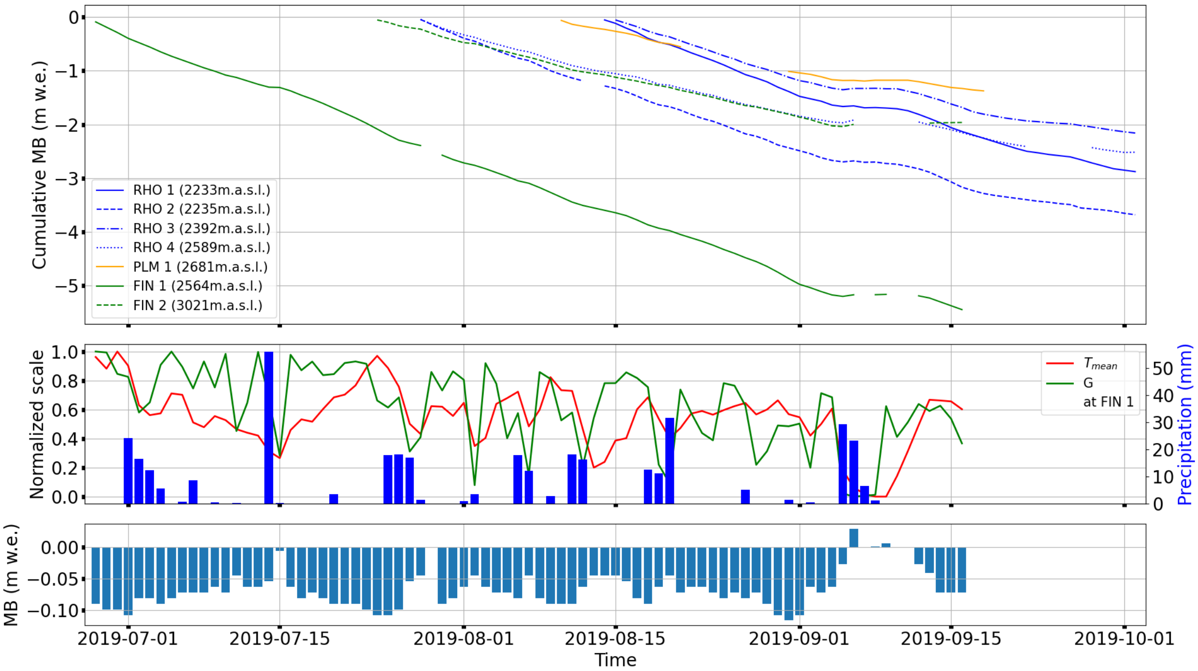
RC: *The title mentions 'mass balance observations', whereas the observations are of surface elevation.*
AR: Our intention was to simplify the wording for the reader. As explained in our response to the comment on the manuscript title by Reviewer #1, we suggest to change the title as follows: "Assimilating near real-time mass balance stake readings into a model ensemble using a particle filter"

RC: *Line 96. It is mentioned that the camera images are read 'manually' to obtain the daily cumulative surface height change. Is it literally reading the marks from the ablation pole? How could this be automated to be applied to more places?*
AR: Yes, we have read the marks manually. We are working on a procedure to automate this though, so that operational runs that we plan for the future won't require manual interventions. In this respect, see our EGU2021 abstract (https://doi.org/10.5194/egusphere-egu21-7663) and our GitHub repository (https://github.com/leosold/TOAST).

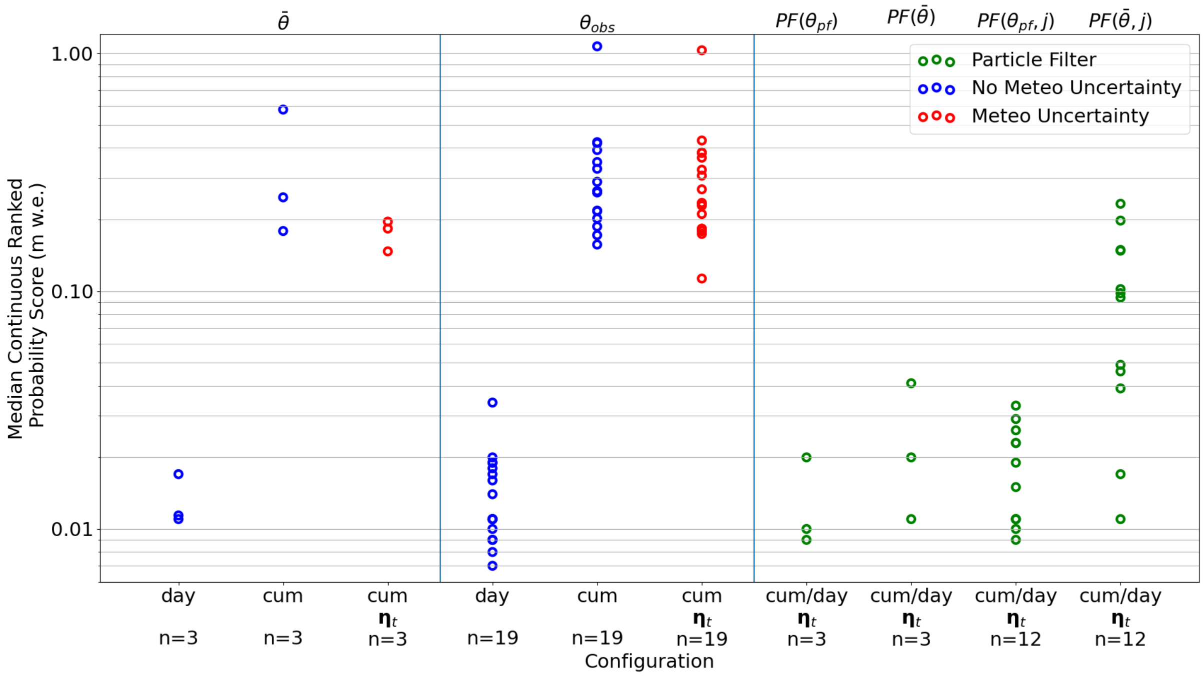
RC: *Line 127. Can you say more about the 0.2 degree resolution? How does this compare with other products? Is it high or low resolution?*
AR: At Swiss latitudes, 0.2 degrees corresponds to a resolution of about 2km. Compared to Global Climate Models or Regional Climate Models, which are sometimes used to force glaciological models

directly, this is a very high resolution. We will specify this in the revised text: "…. which for Switzerland corresponds to a horizontal resolution of about 2 km."

RC: *Figure 7. I think making the vertical axis larger for panels a and b could make the figure easier to read.*
AR: As a response to the request, we suggest to double the extent of these two axes:



RC: *Figure 8. The individual circles are difficult to see. Please make the circumferences thicker, and maybe increase the size of the figure.*
AR: We will increase the line thickness of the circles to improve the visibility and also increase the figure size:



RC: *In pages 22-24 (approximately) there are several places where a quantity is written followed by () or []. It was not clear to me what the quantities in the parenthesis are, and why there are two styles.*
AR: It was explained in l. 371 that we use the square brackets to add the non-proper Continuous Ranked Probability Score (CRPS) into the text, but apparently it needs to be repeated in the Results section. Moreover, the round brackets we use in our notation are to be understood as annotations in the way

they are commonly used. We will add a reminder at the first occurrence of the square brackets: "(proper CRPS outside, non-proper CRPS inside the square brackets)".

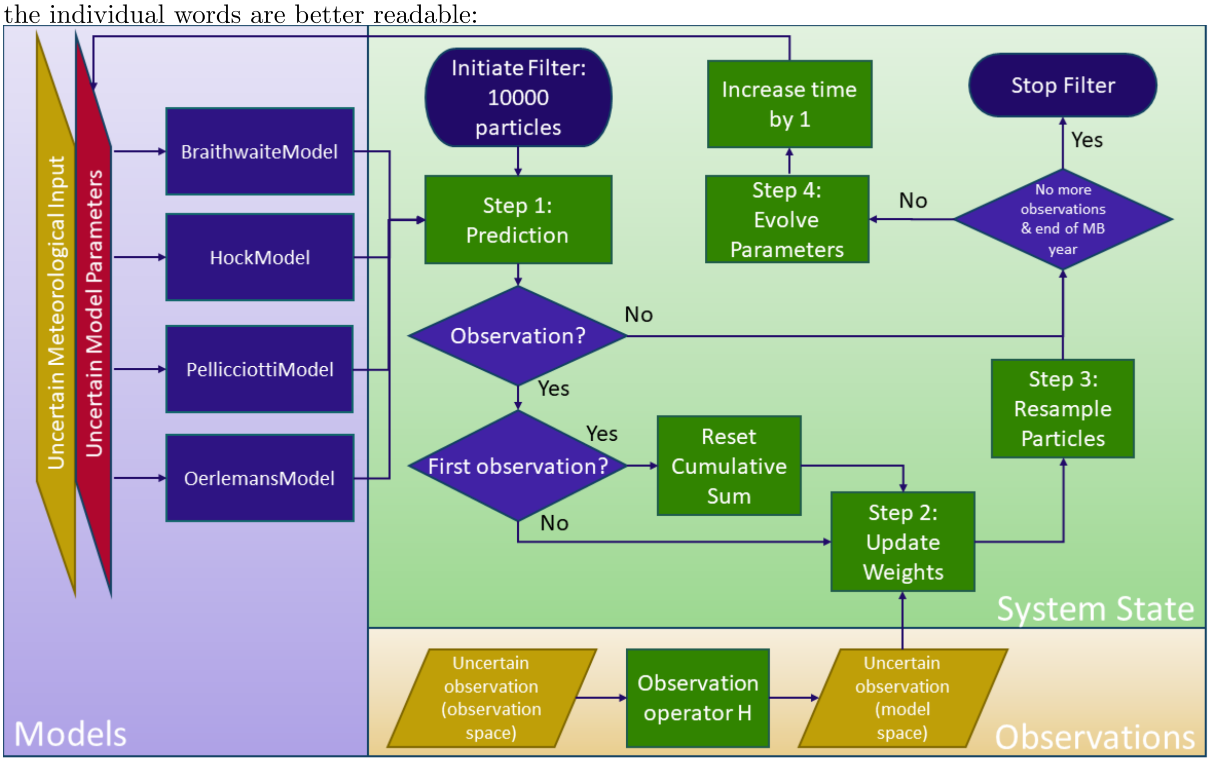
- Typos and corrections -

RC: *Line 109. ... because it can happen that the camera construction sinks...→ ... because the camera construction can sink. (easier to read).*
AR: We will change as suggested.

RC: *Line 118. melt during night → nighttime melting*
AR: We will change this as suggested.

RC: *Line 190 on. When mentioning the models in an itemised list, start the sentences with capital letter.*
AR: We will change this as suggested.

RC: *Figure 5. Some of the words in the labels are split into two lines*
AR: This was a tradeoff between font size and visual appearance. We will reformat the figure so that the individual words are better readable:



RC: *In the title of table 2 it should say 'standard deviations' instead of 'covariances'.*
AR: We thank the reviewer for noticing this. It will be corrected.