General Statement

I am pleased to say that the authors have addressed my major concerns. The Observing System Simulation Experiments setting is now more rigorously set, by using a different (and more complex) version of the model to generate the synthetic truth than the one used in the assimilation ensemble. As expected, this degrades a bit the estimated performance of the assimilation algorithms (the degradation is even very strong for the PF on Fig. 7), but increases the reliability of the statements of this paper.

The inclusion of the synthetic assimilation of Snow Cover Fraction (SCF) is also a nice addition, as it allows to estimate the relative information content of Land Surface Temperature (LST) observations with respect to SCF. As the community has a good understanding of SCF assimilation and information content, presenting LST as a similar, but potentially even more informative variable, is probably the most important statement of this paper.

Please find below some minor comments, the most important ones revolving around the definition of what is the "reference" and how should its performance be measured and asking for a better structure of the data and methods section. I am pleased to provide the authors with an annotated version of the manuscript, containing some comments to address, and some suggestions. I have no doubt that the authors can address these comments easily and am looking forward to seeing this paper published!

Author's response: The authors would like to thank the reviewer Bertrand Cluzet for his positive evaluation. We created another Figure to show the spread of the prior ensemble. We would like to decline the suggestion to add "data assimilation" to the title of our article as it would make the title cumbersome. Otherwise we have followed most of the suggestions as explained below in a point-by-point response.

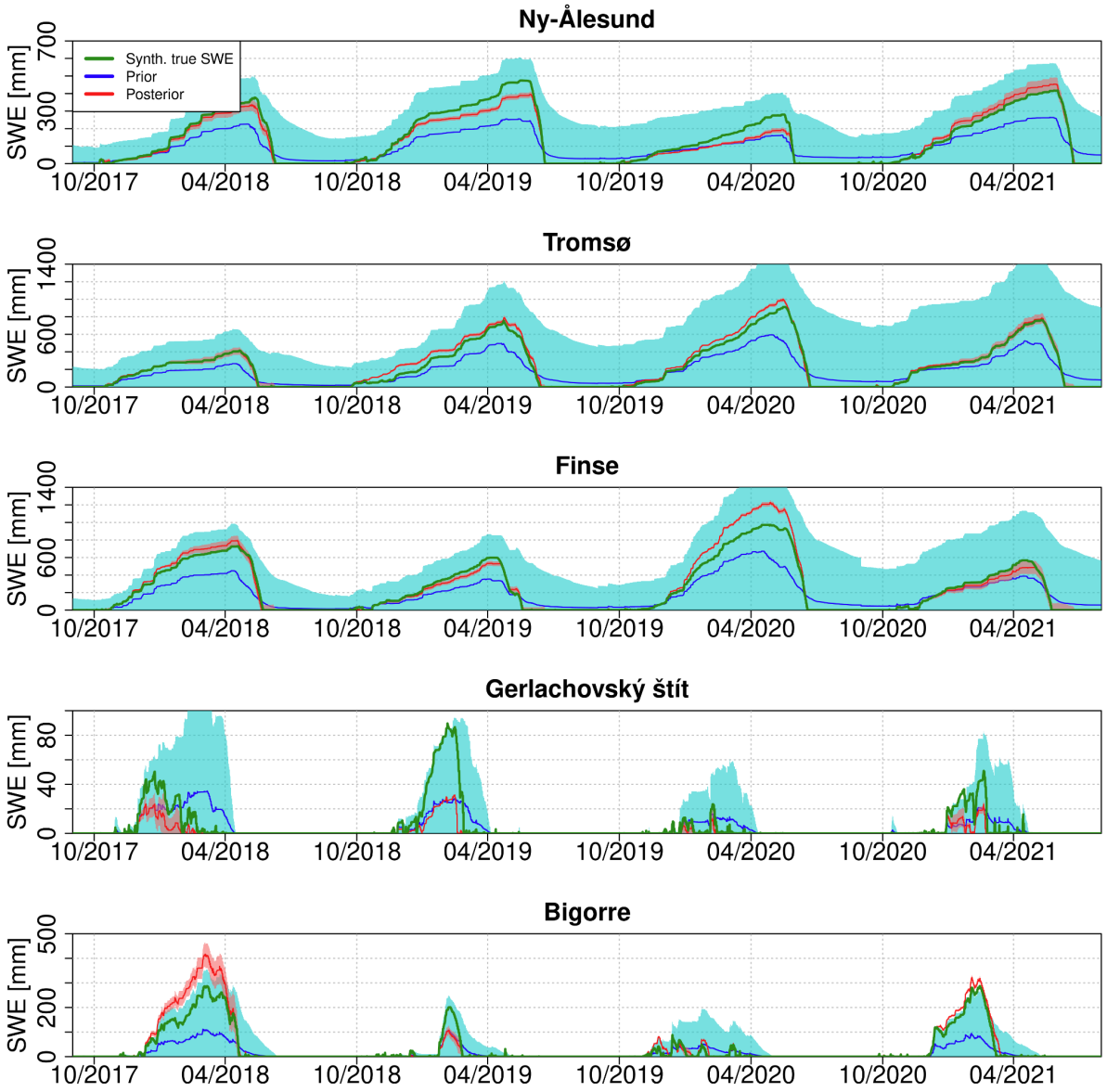
Minor comments:

- The parameters chosen for the log-normal distribution are such (mean= 1.15, median = 1.57) that the prior is likely not centred on the openloop run (which is displayed in the figures, and serves as a reference for the scores), but is already pulled a bit towards the synthetic observations. In other words: the openloop is not representative of the prior, and this could be misleading. From a theoretical standpoint, in order to infer the added information of the data assimilation algorithms, the scores should be computed with respect to the prior mean or median's performance, and not the openloop (see also Aalstad et al., 2018). I would ask the authors to add the score of the prior to all the evaluation plots. Please also add the prior ensemble quartiles (25-75 quantiles + median) to Fig.2 so that the reader can visualize how does the prior look like (as is done in Aalstad et al, 2018, e.g. Fig. 2.). I acknowledge that this point was raised by the last comment of reviewer 2, and rejected by the authors, but I still think that this would be very informative for the reader.

Author's response: We agree that the prior ensemble is useful information to grasp the effect of the assimilation process. However we would like to avoid adding the prior ensemble to Figure 2 for the following reason: Figure 2 already shows ensembles spread for the 3-day and 16-day revisit scenarios, however these ensembles represent the dispersion of the replicates, not of the spread of the posterior ensembles. Therefore, adding an ensemble representing the spread of the prior would bring confusion as there would be two different kind of ensembles in the same figure. Instead, we have created a new figure that represents the prior of a single random replicate for each site and its posterior, for the 3-day revisit

scenario in the absence of clouds. In this figure we show the spread of the prior ensemble as requested by the reviewer using its standard deviation (see figure below).

This figure shows that the prior ensemble is very large and subject to the creation of “glacier” as the excessive snow accumulation or lack of ablation results in a non-zero snow water equivalent over the entire simulation period in three of the study sites, and this bias increases year after year. This makes irrelevant the computation of error metrics, and we believe would lead to an unfair overestimation of the LST assimilation performance. Therefore, we prefer to keep the original strategy of comparing the performance with the deterministic open-loop.



Supplementary 2: Example of prior and posterior SWE ensembles for a single replicate assimilating LST at a time resolution of 3 days in the absence of clouds. The shaded areas show the prior and posterior standard deviation.

- What is the observation error prescribed in the PBS/PF? I could not find this important piece of information in Sec. 2.

Author's response: The errors are the same as those used to degrade the observations. We have clarified this in the text:

"The observation errors prescribed for the synthetic true observations were the same as those defined to generate the degradation Gaussian noise."

- Please substantially improve the writing of the data and methods section, by splitting it into subsections, and adding several important pieces of information (nRMSE is not introduced, observation errors are not specified, model parameterization are only vaguely described)

Author's response: We have splitted this section in 3 subsections. We have also included the observation errors and nRMSE. However, we consider that it is not necessary to add a detailed description of the FSM parameterizations in this paper. We have already included the compilation numbers in table 1, hence the curious reader can easily find the information within the FSM repository and its original publication in GMD.

- MuSA version: the authors say in their response that "there is no specific tag for the version used here, so we just included the version number. The modifications mentioned here are in the FSM code, but is also included in the MuSA version". Sorry to insist, but having a clear tag or commit reference is essential to ensure science reproducibility. I found the following tag on the github repo: https://github.com/ealonsogzl/MuSA/tree/v1.0. If this corresponds to the exact version of the code that was used for the paper, please put it into a Code and Data availability section, otherwise just provide a new tag.

Author's response: We have added the link in the code and data availability section as suggested by the referee:

"The MuSA v1.0 code is available at github.com/ealonsogzl/MuSAhttps://github.com/ealonsogzl/MuSA/tree/v1.0."

- Briefly compare FSCA-DA results with the literature. I think that dividing RMSE by a factor of 2-3 as shown in Fig. 5 is in the ballpark of the literature with real SCF data assimilation (see e.g. Aalstad et al., 2018). This could help convincing the reader that this OSSE setup provides a realistic estimate of the information content of LST and SCF.

Author's response: We thank the referee for this suggestion. We have included the reference suggested by the referee to highlight that the error metrics are comparable to those found by previous studies:

"The results indicate a comparatively lower reduction in SWE nRMSE when FSCA is assimilated compared to LST assimilation ranging between 27% and 39% in the case of the 3-day revisit experiments and 41% and 45% for the 16-day revisit experiments (Figure 5). These results are comparable to those obtained in previous studies assimilating real observations, confirming the robustness of the OSSE design (Aalstad et al., 2018)."

Technical comments:

- Fig. 2: the spread of the different "replicates" is a nice addition. Perhaps consider making it more clear to the reader that here, the replicates differ only from observation error noises (and not from e.g. different input perturbation scenarios), either in the text, or the figure legend.

Author's response: We have added the following sentence in the text to clarify this point:

"*The posterior SWE series in Figure 2 were averaged from an ensemble of 100 replicates under clear sky conditions. The uncertainty in the replicates is caused mostly by the different random noises in the observations (see Sect. 2. Data and methods).*"

- Try to include the keywords 'assimilation' 'data assimilation' in the title.

Author's response: As stated before, we prefer to keep the title in its current form, in order not to increase its complexity.

Annotated comments:

l.23: Please introduce how this score is computed in the data and methods section

Author's response: Added.

l.130: References are needed here to describe the parameterizations. I know that the current version of FSM is not published yet, but you could include a reference for each of the parameterizations, or insert the equations yourself. Otherwise, the terms of Table 2 are too vague for the reader to really understand.

Author's response: As mentioned above, we believe it is unnecessary to add detailed descriptions of the different FSM2 parameterisation schemes.

Table1: ... in particular, more details are needed on this parameterization, as this is your "observation operator" for the SCF assimilation.

Author's response: These two parameterisations are very simple, and their names are self-descriptive. Option 1 is a linear function (fsnow = min(snd/hfsn, 1.), where snd is the SWE and hfsn is the snowcover fraction depth scale) and option 2 is an asymptotic function (fsnow = snd / (snd + hfsn)). We do not consider further explanations necessary

l.146: I just realized that with these parameters, the "openloop ensemble" is not centered on the openloop (mean = 1.15, median = 1.57), which artificially pulls the prior in the good direction.

Author's response: In reply to a previous comment, we have added a plot representing the dispersion of the prior, so this is also answered.

l.150: PF

Author's response: Corrected.

l.168: ... applied to the observations

Author's response: Not both noises are applied to observations, but we will recall this information here as suggested:

"Also the position of the gaps in the different cloud cover scenarios and the Gaussian and Brownian noises added to the observations and forcing respectively are random"

l.169: observation noise

Author's response: The Brownian noise is not applied to the observations.

l.184: figure

Author's response: Corrected.

l.184: This formulation is misleading. I think that you are talking about the spread of the mean of the replicates (each replicate is a posterior ensemble in itself), and not the spread of each replicate (what your formulation tends to mean)

Author's response: Corrected.

l.203: were

Author's response: Corrected.

l.237: The reader has to believe you on this interpretation: this is impossible to tell from the aggregated scores. Please provide timeseries to substantiate this statement, or consider removing it.

Author's response: We have included a reference to the new figure that shows the creation of glaciers.

l.268: I don't think that this is arguably a ' mitigation', this is just the standard procedure for OSSE's. I understand your point, but please consider using a different formulation

Author's response: Added: *'as typically done in OSSE's'.*

l.274:Please be more specific, whether you mean intra pixel variability of snow conditions (then Robledano etl. al., 2022 could be cited) or just mixed pixel (snow/ground/rocks...)

Author's response: Both effects could be considered depending on the situation, we have add this explanation in the text:

"Also, the surface temperature observation in complex terrain may differ from the simulated temperature, due to intra-pixel variability as a consequence of variable snow and/or mixed pixel conditions (Robledano et. al., 2022, Lundquist et al., 2018)."

l.288: ... showing that the performance of the assimilation would depend on cloud cover scenarios (suggestion)

Author's response: Accepted