Review of Oberrauch et al. (2026) 10.5194/egusphere-2026-2725

General comments

This manuscript by Oberrauch et al. presents a novel and promising approach to operational snow data assimilation that allows for the seamless incorporation of satellite-based fractional snow-covered area (FSCA) retrievals from Sentinel-2 into the existing FSM2OSHD framework via a particle batch smoother (PBS). The proposed method is a two-step approach. The first step uses an established method for assimilating sparse in-situ snow depth observations in flat terrain to update atmospheric forcing perturbation parameters via a particle method previously described in Oberrauch et al. (2024). The novel second step uses the PBS to assimilate FSCA retrievals to update two annually static parameters controlling albedo decay and a precipitation multiplier in sloping terrain. Such a two-step approach is made possible by the clever move of restricting the FSCA updates to sloping terrain, such that the two assimilation steps are effectively independent of one another. As such, the two assimilation steps can be performed sequentially in an online manner, making this approach well suited for operational snow hydrological forecasting applications. The proposed method is tested for the entire, primarily Swiss, FSM2OSHD operational domain, with a particular focus on the Lake Thun and Inn catchments. Tests are performed through four experiments: 1) a control experiment with station-based snow depth assimilation only, 2) the two-step approach including FSCA assimilation applied at the pixel level, 3) applying a subsequent spatial interpolation step for the sloping terrain parameters, 4) a 'leave-one-pixel-out' approach that interpolates from neighboring pixels but ignores the local pixel-level parameters. These experiments demonstrate a substantial improvement in simulated FSCA performance from the two-step approach for observed pixels, particularly during the ablation season.

Overall, the experiments in this study are mostly convincing, and the manuscript is generally well written. Most importantly, this is a timely and highly valuable contribution to the snow data assimilation literature that demonstrates, possibly for the first time, how high-resolution satellite data from Sentinel-2 can be incorporated into an operational snow hydrological forecasting workflow in the form of FSM2OSHD. The use of a two-step approach that decouples flat from sloping terrain is particularly innovative and serves as a clever method to keep the added cost of running a PBS for FSCA data assimilation tractable. It is very encouraging to see satellite data being used in operational snow data assimilation workflows, which has historically been a largely untapped resource despite the widespread use of FSCA for snow reanalysis and the fact that Sentinel-2 has been in space for over a decade.

Reading through the manuscript prompted three larger questions in my mind, though. Firstly, why did you restrict validation to only using FSCA (SCF) data from Sentinel-2? There was a missed opportunity to check if FSCA assimilation improves spatially distributed estimates of snow depth using airborne data such as that in Magnusson et al. (2025). Such an extended validation might be beyond the scope of this work and rather the topic of another study, but many of the results in this manuscript are (for now at least) restricted to improved performance in terms of estimating FSCA as opposed to snow depth or SWE. Secondly, the choice of either doing pixel-by-pixel updates or full interpolation seems like a false dichotomy. Why not do something in-between where one only interpolates the PBS target parameters for unobserved (by Sentinel-2) grid cells but keeps the local effects captured at observed locations? Thirdly, I was surprised that the precipitation multiplier should be bounded from above by 1. This may make sense in terms of parametrizing gravitational redistribution on slopes, but in terms of correcting biases in snowfall in unobserved locations I would expect multipliers to potentially be well above 1. In particular, the snow depth stations used in the first pass to correct subregional forcing errors surely do not cover the most extreme high local snowfall conditions in the Swiss Alps? I return to these and other points in the specific comments below.

In summary, this manuscript was a pleasure to read and is clearly well suited for publication in *The Cryosphere*. I recommend publication after minor revisions to address the questions above and the specific comments below.

Specific comments

1. Abstract: *"... assimilation of point-scale snow depth observations accounts for spatiotemporal uncertainties in the meteorological forcing at subregional scales"*, this claim is surely based on Oberrauch et al. (2024), but it seems hard to reconcile point-scale snow depth observations with a (presumably much larger?) "subregional"

scale. To me, this seems to presume that these stations are all located in "subregionally" representative locations, is that really always the case? Consider rewording. I am not questioning this method, but rather reacting to the scale-mismatch.

2. L58: Very nice introduction. Change *"via Ensemble Smoother techniques"* to *"via batch smoother techniques"*. The Ensemble Smoother (ES) is a particular technique, namely the batch smoother variant of the Ensemble Kalman method that is distinct from other Ensemble Kalman Smoothers (Evensen et al., 2022). Batch (also known as fixed-interval) smoothers are the Bayesian approach under which both ES and PBS fall.

3. L107: Perhaps a slight or at least apparent inconsistency here, what was previously referred to as gravitational redistribution is now *"precipitation input"*. The latter is presumably a way to parametrize the former? Consider clarifying or maintaining a more consistent terminology throughout. Also, what unresolved process is being captured by varying the albedo decay parameter at the grid cell level, is it variations in light-absorbing impurities or something else related to shortwave radiation?

4. L123: Avoid defining a symbol for snowfall via SF in LaTeX, mathematically this reads as if you are multiplying two variables S and F. The standard approach is either to stick to single letter symbols in the usual math font and use roman symbols for abbreviations in the form SF (i.e., `\mathrm{SF}` in LaTeX). Either change would conform to the standard for Copernicus Publications that follows the math typesetting regulations given by the IUPAC Green Book (see the `template.tex` Copernicus template file).

5. L145: Same here in Eq. 3, define the precipitation multiplier either as `\mathrm{pm}` or use a single letter.

6. L157: Same here in Eq. 4, define the snow cover fraction as `\mathrm{SCF}` or use a single letter.

7. L174: Change *"which reflects the probability of observing the given value under the assumption that the particle represents the true state of the system"* to *"whose value is evaluated numerically as the probability density of the fixed observation, given the assumption that the particle represents the true state of the system."*. The likelihood function can be a confusing concept, but this second formulation is more correct (the likelihood is not a probability) and in line with the likelihood principle (see e.g. MacKay, 2003). The likelihood is a function of the uncertain model state and/or parameters rather than the fixed observations and can be effectively seen as a relative measure of goodness-of-fit between the model and observations.

8. L176: Change *"filter degeneracy"* to *"weight degeneracy"*. Resampling does not actually mitigate filter degeneracy, it just mitigates weight degeneracy. Typically, the filter tends to collapse as there is still path degeneracy (all trajectories are the same) unless some form of particle rejuvenation is applied. In practice, it is important to emphasize this distinction as path degeneracy matters much more than weight degeneracy (see e.g. Murphy, 2023). For the snow depth assimilation step you ensure particle diversity by reverting to the prior parameters for each window, which from what I recall serves as a form of rejuvenation or jitter in your case.

9. L180: Change *"absence of any Gaussian assumptions"* to *"absence of the linear assumption"*. What makes the particle filter well suited here is not the lack of a Gaussian assumption; you use a Gaussian likelihood, and many of your priors are (up to a transformation) Gaussian. Rather, it is the lack of a linearity assumption. The fact that particle methods do not need Gaussian assumptions is a nice feature, but you (and the rest of us in snow science) are not yet leveraging that. Either way, at least for now, nonlinearity is likely the bigger benefit over ensemble Kalman methods for snow data assimilation.

10. L184: Change *"evidence"* to *"information"* since evidence (marginal likelihood) means something specific in the context of Bayesian methods. So does information, but it arguably fits better here.

11. L187: Change *"other ensemble smoothers"* to *"other batch smoothers"*.

12. L194: You are presumably perturbing these independently for each 3-day window by resampling from the prior. If so, please specify that so that the reader does not need to read Oberrauch et al. (2024) to learn this.

13. L203: Eq. (5) please change this to the correct formulation of the Gaussian likelihood, namely:

$$\omega_i^t \propto \exp\left(-0.5\left(\frac{d_{\mathrm{obs}}^t - d_{\mathrm{sim},i}^t}{\sigma_i^t}\right)^2\right). \tag{5}$$

Even if the ordering $d^t_{\text{obs}} - d^t_{\text{sim},i}$ or $d^t_{\text{sim},i} - d^t_{\text{obs}}$ does not matter numerically in this case, the current version is strictly wrong and invites a backward interpretation of the likelihood. As you said nicely around L174, the likelihood operates under the assumption that the particle represents the true state of the system. In that case, with a Gaussian likelihood, $d^t_{\text{sim},i}$ plays the role of the mean and the residual is defined as $\epsilon^t_i = d^t_{\text{obs}} - d^t_{\text{sim},i}$ which corresponds to the innovation in (ensemble) Kalman methods. Getting this order right will be increasingly important when the community moves beyond Gaussian likelihoods to leverage the full potential of particle methods. As an even more pedantic side note, I would discourage using t as a time index since t is usually reserved for the timestamp itself. Consider instead using k to index time t_k. Also, using the time index as a subscript and the particle index as a superscript is arguably more elegant, i.e., $\omega^{(i)}_k$.

14. L204: Change *"observation uncertainty"* to *"observation error standard deviation"*. Uncertainty is a broader concept than Gaussian standard deviation.

15. L207: For someone not familiar with Oberrauch et al. (2024), collapsing deliberately seems like an odd choice. Consider flagging already here that this is done to facilitate parameter interpolation.

16. L233: Why did you pick a factorial combination of parameter settings on a regular grid in parameter space as opposed to actually sampling from a prior? The latter would be more consistent with the PBS, while what you are doing now is closer to some kind of quasi Monte Carlo method. There is nothing necessarily wrong with this, it is just not motivated well here.

17. L240: Eq. (6), nice here you got the order in the exponent right. Maybe Eq. (5) was just sloppy.

18. L243: I don't think van Leeuwen said this explicitly in 2012 unless he time traveled back from when the PBS paper was published (Margulis et al., 2015). He probably said something more general like SIS is SIR without resampling. Regardless, consider also citing here Alonso-González et al. (2022) that actually points this out directly when it comes to the PBS.

19. L256: Not following why cloud cover leads to unobserved pixels with the PBS when the window is the whole water year, unless a pixel is permanently cloud covered or you use just a few Sentinel-2 scenes.

20. L259: The choice of smoothing observed pixels seems odd, why not just interpolate only for the unobserved pixels and keep the observed pixel parameters at their local unsmoothed values?

21. L262: Change the SW notation unless you are multiplying S and W.

22. L268: Not following why the albedo decay rate should be scaled by 20^{-1} when $S^{\downarrow}_{\text{inc}} = 0$ but when both are 0 it is scaled by $20^0 = 1$? There is likely a good reason for this given that the authors emphasize this point, please clarify.

23. L278: Using the symbol z for this is unfortunate as it is often used for elevation. Maybe this was just a typo. Why not just write $\zeta = 37°$ here?

24. Figure 3: This parametrization seems to limit the multiplier to the range $0 < \text{pm} \leq 1$ so FSCA data is unable to increase the solid precipitation beyond the nominal multiplier identified in step 1. This might be fine from a gravitational redistribution in sloping terrain point of view, but I would expect there to be locations at elevation that could have a precipitation multiplier greater than one reflecting biases in the snowfall forcing.

25. L304: Change *"multi-spectral"* to *"multispectral"* since this is referring to unmixing multispectral (as opposed to say hyperspectral) satellite imagery.

26. L310: Not sure how masking for clouds mitigates these issues? Given that discriminating clouds from snow can be challenging, don't cloud masks inherit exactly these issues? Or do you mean manual cloud masking?

27. L320: Here it is surprising that a mix of experiment 2 (pixel-by-pixel) and 3 (interpolation) was not attempted where you interpolate for only the unobserved pixels. That would get rid of some of the undesirable smoothing of local features, and would presumably not be any harder to implement than experiment 3 or 4.

28. L321: An alternative and arguably at least as informative approach to fully independent validation would have been to use some independent SWE or snow depth data that has not been assimilated, such as from airborne campaigns.

29. L335: As part of your data analysis for the results, did you check FSCA errors specifically for flat terrain where it is assumed that no additional albedo decay or precipitation multiplier correction is required? If there are large errors remaining after step 1, there is presumably something to be gained from FSCA assimilation here too. Part of the elegance of your scheme is a decoupling between flat and sloping terrain, but it would still be worthwhile diagnosing if you really don't need any correction for flat terrain. Especially for the flat pixels that do not contain snow depth observations.

30. L345: Strictly speaking, this is only demonstrated in terms of matching FSCA (SCF) not SWE or snow depth. Your past work Oberrauch et al. (2024) has shown convincingly that this step 1 helps at observed (and unobserved) flat station locations, but it remains to be shown explicitly that this actually helps improve SWE estimates at the subregional scale. So what you are saying here is not necessarily wrong, but perhaps too strong of a claim given what you show in Figure 4. Consider softening or specifying that this is in terms of FSCA. The reasoning to make this distinction is that inferring SWE from FSCA notoriously suffers from non-identifiability (equifinality), in that you can reach the same FSCA by having accumulation compensate for ablation and vice versa. So just because you match FSCA at the subregional scale it does not necessarily follow that you match SWE.

31. L352: As for L345, this is only shown in terms of FSCA which is a necessary but not sufficient condition for demonstrating improved estimates of other snowpack states (SWE, snow depth).

32. Figure 4: Since, unlike the likelihood function that treats the simulated state as true, you are taking the observed Sentinel-2 SCF map as a reference here, wouldn't it make more sense to plot the simulated SCF error which is Sim-Obs not Obs-Sim?

33. Figure 6: On a related note, the simulated SCF error (not bias, which would be some average of this error) should be defined as Sim-Obs not Obs-Sim when you are treating the observations (Obs) as the reference truth.

34. L404: In light of the previous comment, check that you have defined bias properly as the average of Sim-Obs for each observed date. Comparing Figures 7 and 6, it looks like your bias definition in Figure 7 is correct (mean of Sim-Obs) since you have a mostly positive bias (overestimation of SCF) in Figure 6 for 04.05.2023 when accounting for the fact that you plot Obs-Sim there.

35. L436: Please be more specific here too (see comment on L345), saying that you improve distributed snowpack estimates is a bit too strong a claim since you have only compared to SCF data. As such, change *"snowpack"* to *"snow-covered area"*.

36. L458: Nice discussion! Another idea worth mentioning would be to use the SCF data to infer the "background" parameter climatology that could help to have more informed locally tailored priors on the albedo decay and precipitation multipliers for each new water year. These parameters might be relatively consistent from year to year for each pixel as shown in e.g. von Kaenel and Margulis (2025).

37. L465: Change *"exist"* to *"be informative"*. This is a seemingly small, but nonetheless important semantic change. There is a distinction between the spatial correlation in a physical field and the spatial correlation in the probability distribution of model parameters. It is the latter we are interested in for data assimilation, and I agree that we can certainly model it in terms of geographic and topographic parameters. Still, I don't think it makes sense to speak of correlations on the distribution of model parameters, related to epistemic uncertainty, as existing in the same way as a snow depth field does.

38. L483: When discussing structural uncertainty related to the choice of SCF parametrization, consider citing Sourp et al. (2026) who demonstrated the importance of this when using the PBS for snow reanalysis.

I would like to congratulate Oberrauch et al. on this innovative study that will undoubtedly help the community make satellite-based snow data assimilation more widely applicable, especially in operational contexts.
Kind regards,
Kristoffer Aalstad

References

Alonso-González, E., Aalstad, K., Baba, M. W., Revuelto, J., López-Moreno, J. I., Fiddes, J., Essery, R., and Gascoin, S.: The Multiple Snow Data Assimilation System (MuSA v1.0), Geosci. Model Dev., https://doi.org/10.5194/gmd-15-9127-2022, 2022.

Evensen, G., Vossepoel, F. C., and van Leeuwen, P. J.: Data Assimilation Fundamentals, Springer, https://doi.org/10.1007/978-3-030-96709-3, 2022.

MacKay, D. J. C.: Information Theory, Inference, and Learning Algorithms, Cambridge University Press, 2003.

Magnusson, J., Bühler, Y., Quéno, L., Cluzet, B., Mazzotti, G., Webster, C., Mott, R., and Jonas, T.: High-resolution hydrometeorological and snow data for the Dischma catchment in Switzerland, Earth System Science Data, 17, 703–717, https://doi.org/10.5194/essd-17-703-2025, 2025.

Margulis, S. A., Girotto, M., Cortés, G., and Durand, M.: A particle batch smoother approach to snow water equivalent estimation, Journal of Hydrometeorology, 16, 1752–1772, https://doi.org/10.1175/JHM-D-14-0177.1, 2015.

Murphy, K.: Probabilistic Machine Learning: Advanced Topics, MIT Press, URL `http://probml.github.io/book2`, 2023.

Oberrauch, M., Cluzet, B., Magnusson, J., and Jonas, T.: Improving Fully Distributed Snowpack Simulations by Mapping Perturbations of Meteorological Forcings Inferred From Particle Filter Assimilation of Snow Monitoring Data, Water Resources Research, 60, https://doi.org/10.1029/2023WR036994, 2024.

Sourp, L., Pedinotti, V., Alonso-González, E., Jarlan, L., and Gascoin, S.: Assessment of Snow Cover Fraction Parameterizations for High Resolution Snowpack Reanalyses, Hydrological Processes, https://doi.org/10.1002/hyp.70491, 2026.

von Kaenel, M. and Margulis, S. A.: Improved modelling of mountain snowpacks with spatially distributed precipitation bias correction derived from historical reanalysis, The Cryosphere, 19, 3309–3327, https://doi.org/10.5194/tc-19-3309-2025, 2025.