

Summary

Singh et al. extend the physical-statistical Bayesian SWE retrieval framework from 2024 work by the same lead author from open areas to forested landscapes, using dual-frequency (X- and Ku-band) SnowSAR backscatter observations from the SnowEx'17 campaign across Grand Mesa, Colorado. The authors couple a multilayer snow hydrology model (MSHM) with the Microwave Emission Model of Layered Snowpacks (MEMLS) and introduce a modified Water Cloud Model (WCM) to represent canopy backscatter and transmissivity, adding an intermediate Bayesian step to estimate vegetation parameters (A , B , M_v) alongside snow and ground properties. Retrievals at 90 m and 30 m resolution are evaluated against snow pit SWE and ASO LiDAR snow depth observations, with reported RMSE of 0.033 m (~8% of maximum pit SWE) and improved agreement with LiDAR relative to MSHM priors alone. Performance is strongest for pixels with forest fraction greater than 70% and at finer (30 m) resolution, where mixed land-cover contamination is reduced; error is concentrated at forest edges, ecotones, and in topographically complex terrain.

This is a meaningful, timely, and reproducible contribution to forested SWE retrieval, outlined to meet the NASEM Decadal Survey requirements. The figures are clear throughout, and the conceptual ones contribute significantly to outlining the retrieval workflow.

Primary feedback centers on quantifying uncertainty that the manuscript currently treats qualitatively or not at all.

Comments

1. The transfer of ground parameters from open areas to forested pixels (Section 3.2.2, lines 230–238) currently depends on low spatial variability among open-area estimates alone, as stated in the text — but this does not establish that the same values hold under canopy. A sensitivity test perturbing the interpolated ground parameters to see how much the final SWE retrieval changes would help address this, given that in-situ validation data under canopy may not be available.
2. The WCM parameter-fixing scheme (Section 3.2.3, line 254; discussed further via retrieved A/B in Appendix D) is presented without addressing how sensitive retrieval skill is to which parameter is held constant at each step.
3. The 30% RRB threshold used to define a “successful” retrieval (Section 3.3, line 295; Conclusion, line 409) is described as chosen “as a balance between physical quality and data retention” rather than from an independent physical or statistical justification — it would help to know why 30% specifically, rather than, say, 20% or 40%. Pixels that fail this threshold (along with the 30°–50° incidence angle criterion) are excluded from the reported skill statistics (e.g., the ARD percentages, Figures 6/8/9) — please report what fraction of all forested pixels are excluded; to help readers judge how representative the reported statistics are of overall retrieval performance.
4. The snow pit validation (Table 7, lines 315–323) pools pits with only one collocated retrieved pixel within 100 m at equal weight with pits that have many more (e.g., $N=14$) — a weighted or stratified error metric, or a check on how the reported 0.033 m RMSE

changes if the single-pixel pits are excluded, would clarify how much those pits are driving the result.

5. The fourth flight's poor performance (Table 8; lines 348–354, 369–377) is attributed to retrieval error, but the authors' own citation of LiDAR underestimation under dense canopy and steep terrain (Jacobs et al., 2021) is never used to separate how much of the disagreement is LiDAR bias instead.
6. Canopy closure and the interception model were both developed at other sites and applied to Grand Mesa without site-specific validation. The manuscript states that uncertainty from these inputs has not been propagated into the final retrievals (line 428, line 484). Short of a full perturbation-and-rerun test, the authors could instead (a) discuss, even qualitatively, the expected direction and magnitude of error these relationships might introduce given known differences between Grand Mesa and the sites where they were originally developed, or (b) report any uncertainty bounds already published alongside the original Pomeroy et al. (2002) and Hedstrom and Pomeroy (1998) parameterizations, to give readers some sense of the error these inputs could be carrying forward.

Line-by-line feedback

- Line 75–77: The reported RMSE (<7%) is presented as a favorable result — can the authors clarify how this threshold was established, e.g., whether it reflects accuracy standards accepted elsewhere in the snow remote-sensing literature (e.g., reiterate NASEM Decadal Survey requirements)?
- Line 109: Which HRRR forecast hour was used? Given that forcing data are subsequently modified within the workflow, it would help to note whether forcing-data limitations are addressed later in the manuscript.
- Line 126: A reference is needed here.
- Line 127–128: Please cite the prior analyses referenced here.
- Table 1 / Line 147: How many snow pits were evaluated, and on what basis were these specific selected? This information (number of sites) appears in Table 7 but is not stated in the main text.
- Line 167–168: Please specify how the snow hydrology model was modified to simulate snowpack physics under canopy, or reference where this modification is described.
- Figure 4: Please clarify what the two columns containing "Veg," "Grd," and "n" outside of the IV steps represent.