Summary and overarching comments:

This study evaluates the impacts of an alternate SCF parameterization, that calculates SCF as a function of topographic complexity and SWE, in offline 1-degree CLASSIC model simulations. The alternate SCF parameterization tends to improve the accuracy of SCF simulated in CLASSIC during winter months in topographically complex regions. This study also explores the robustness of results to differing metrological forcing sources, which reveals the large impact of metrological forcing in snow simulation accuracy. This study provides a novel and important advancement for the CLASSIC modeling system that seems to allow land model simulations to better capture SCF, and in turn improve land-atmosphere interactions due to the snow-albedo feedback. Overall, the paper is well written and the study will warrant a publication after addressing the comments below.

I have four overarching critiques for this analysis. (1) A key motivation for improving SCF in model simulations is to enhance simulated albedo. Although the study briefly covers the impacts of SCF on albedo accuracy using the AMBER score, it would be useful to go into more detail on the albedo analysis which is a critical component of this study. (2) MODIS SCF has questionable accuracy, particularly for representing ground SCF which the land model simulates. This point should be more directly addressed with the consideration of other data sources for SCF (e.g., STC-MODSCAG across the western US, see suggestion below). (3) Discrepancies in spatial resolution between reference data used to validate model simulations and the spatial resolution of the model simulations can largely impact results. Please see specific comment below addressing this point. (4) Figure quality should be improved throughout.

Specific recommendations:

Paragraph starting in line 81: Note that some land surface models also consider SCF as a function of snow density and land cover classification (e.g., He et al., 2023).

He, C., et al. *The community Noah-MP land surface modeling system technical description version 5.0*. NCAR Technical Note NCAR/TN-575+ STR, doi: 10.5065/ew8g-yr95, 2023.

Section 2.1: please articulate the capacities in which CLASSIC is used, for either research applications or operational modeling.

Section 3.1 and Figure 1: Please add information on the calculation of topographic standard deviation. Specifically, what is the resolution of the elevation product which is used to calculate this metric?

Section 3.2: Another potential issue with MODIS SCF is not just its accuracy, but also whether its retrieval represents pixel scale SCF or just the ground SCF. Many land models simulate ground SCF, rather than total pixel SCF (e.g., including vegetated fractions of the pixel) and thus a comparison with the MODIS data used here may not be appropriate. The STC-MODSCAG data addresses this issue, and the latest version has available data across the mountainous western US (https://nsidc.org/data/stc_modscgdrf_hist/versions/1#anchor-data-access-tools). Please consider using these data as an additional reference to evaluate whether the comparisons against MODIS are reliable.

Lines 315-316: Simulated snow density is also a source of SCF uncertainty, e.g., Abolafia-Rosenzweig et al. (2024), which could be noted here or in the Discussion.

Abolafia-Rosenzweig, Ronnie, et al. "Evaluating and enhancing snow compaction process in the Noah-MP land surface model." *Journal of Advances in Modeling Earth Systems* 16.2 (2024): e2023MS003869.

Section 3.3: These SWE evaluations are likely largely impacted by discrepancies between observed and modelled spatial resolutions. It would be good to emphasize this point further, even in the case of airborne gamma SWE observations. To consider the spatial representativeness of observations, consider comparing time series from in-situ stations contained by the same 1-degree pixel and consider whether there are large discrepancies (e.g., with bias and correlation metrics).

Also, when observations are measured infrequently (e.g., a few times in a month) are the modelled data screened temporally to match the observational frequency prior to comparison?

Line 405: Is there truly no feedback in these offline runs? Land models often calculate 2-m air temperature prognostically which could impact SWE. If this is the case for the CLASSIC model, consider re-wording here.

Lines 424-430: Adding more quantitative information here would be useful.

It looks like the simulations tend to underestimate SWE substantially; however, there is a tendency to overestimate winter SCF, largely in the control simulations and modestly in the SL12 simulations. If the SCF scheme is truly accurate at converting SWE or snow depth to SCF then

we would expect to see underestimates in SCF. Can this point be added, particularly connecting logic between Sections 4.2 and 4.3?

It would be interesting to consider whether there are significant correlations between SCF biases with topographic complexity in each of the simulations, and in particular highlight if the SL12 scheme reduces or removes this relationship.

Section 4.4: It would be valuable to note whether the albedo biases are consistent with SCF biases (e.g., locations with SCF overestimates have albedo overestimates).

Section 5: here are some potentially useful references for land model SWE biases:

He, Cenlin, et al. "What causes the unobserved early-spring snowpack ablation in convection-permitting WRF modeling over Utah Mountains?." *Journal of Geophysical Research: Atmospheres* 126.22 (2021): e2021JD035284.

Abolafia-Rosenzweig, Ronnie, et al. "Implementation and evaluation of a unified turbulence parameterization throughout the canopy and roughness sublayer in Noah-MP snow simulations." *Journal of Advances in Modeling Earth Systems* 13.11 (2021): e2021MS002665.

Chen, Fei, et al. "Modeling seasonal snowpack evolution in the complex terrain and forested Colorado Headwaters region: A model intercomparison study." *Journal of Geophysical Research: Atmospheres* 119.24 (2014): 13-795.

von Kaenel, Manon, and Steven A. Margulis. "Evaluation of Noah-MP snow simulation across site conditions in the western United States." *Journal of Hydrometeorology* 25.9 (2024): 1389-1406.