Review of “Accurate assessment of surface mass balance based on stake observations and model correction at Dome A, East Antarctica”

This manuscript presents a substantial modeling effort combining a firn densification model with stake observations to estimate surface mass balance at Dome A. The methodological work on the firn densification model is thorough. The primary conclusion is that firn densification is a major source of error in stake-derived SMB estimates, which can be addressed by using a firn densification model. The paper’s title would suggest a stronger focus on the actual SMB result, thus, I suggest to contextualize the derived SMB value.

Major Comments

[1] While Sect. 3.1–3.2 provide a thorough characterization of the FDM (density profile calibration, column-wide subsidence), the step that most directly supports the paper's title, the subsidence at the stake bottom depth and its use to correct stake-derived SMB (Sect. 3.3), is comparatively brief, and the actual corrected mean SMB values for the two periods (2008–2013 and 2013–2024) are neither explicitly reported anywhere in the text, nor in Fig. 5. Given that these two values are part of the central result implied by the title, I would recommend explicitly stating them (e.g., as annotations and maybe even in Fig. 5), along with an uncertainty range. Since the number of valid stake readings differs between the two periods (and other uncertainties including a mobile stake bottom relative to the surrounding snow layer, strong microrelief around the stake, parallax error arise from the stake method), an uncertainty estimate that reflects these uncertainties would strengthen confidence in these derived SMB values. Additionally, the reported mean SMB of 24.14 kg m^{-2} yr^{-1} is given to two decimal places, implying a high precision. Without a stated uncertainty, this level of precision is difficult to justify.

[2] The discussion of the mean SMB value of 24.14 kg m^{-2} yr^{-1} would benefit from contextualizing against independent observations beyond the RACMO comparison, either from Dome A itself or other regional studies from different sources.

Minor Comments

- Z.61f: This passage remains unspecific in two places. For the Takahashi and Kameda (2007) study: which two methods are being compared? For the Howat (2022) study, please report the percentage rather than (or in addition to) the

absolute number, and specify the measured variable: is it SMB, or only height change?

- Z.91: Please add: “X % of stake height records were missing across the study period and were excluded”
- Z.99: Please add: “X % of AWS data were missing across the study period”. Additionally, the gap filling method could be specified.
- Z. 105ff: The final sentence appears to suggest a level of validation for the use of ERA5 at Dome A in firn densification modelling that is not actually supported by the cited references. My suggestion for rephrasing: “*ERA5 reproduces the variability of near-surface air temperature reasonably well at Antarctic sites where AWS data are available, and, despite uncertainties in precipitation, provides a continuous and internally consistent climate dataset. We therefore use ERA5 Ts and TP as pragmatic driving fields for our firn densification simulations, acknowledging that a dedicated validation at Dome A and for surface (skin) temperature is still lacking.*”
- Fig. 1b) Please mark the ERA5 grid point used.
- Section 2.4.1: Clarification needed for the stake-derived SMB assessment:
    - What is the typical stake depth at Dome A: how much protrudes above the surface, and how much is buried? Please also specify the stake material (cf. Eisen et al., 2008, *Reviews of Geophysics*, Ground-based measurements of spatial and temporal variability of snow accumulation in East Antarctica).
    - Please clarify already in the Methods that, for the densification correction of stake-derived SMB, only compaction down to the buried stake depth is relevant. This would help clarify that the last two terms of Eq. (1) (V_{ice}, dB/dt) are not merely negligibly small, as currently stated for dB/dt, but structurally irrelevant to the stake-based method regardless of their magnitude, since they describe processes well below the depth range the stake actually integrates over.
- Z. 142: Please be consistent with naming: either use "firn densification vertical velocity" as introduced above, or repeat the variable name V_fc here as well.
- Z. 211ff: Does Fig. 2b show two snapshots (one from 2008, one from 2024) of the modelled density profiles, or the annually averaged density profile? Please also explain the purple line (Δρ) in more detail, e.g., that at a given depth z, the same snow layer/isochrone is no longer being compared, since the 2008 and 2024 profiles represent firn of different ages at equivalent depths.
- Z. 233ff: "As the snow compacts, the 2008 surface is buried beneath the 2024 surface, while the stake descends with the surrounding snow layers. Consequently, this cumulative subsidence cannot be used for stake correction." Please clarify: this is because the cumulative subsidence applies to the whole column, not to the specific portion of the column that the stake actually "sees."

- Z. 256ff: For comparison, you could also provide the corrected SMB value and bias estimate for stake depths of 0.8 m and 3 m.
- Z. 269f: Did you also density-correct the SMB for the sub-periods 2008-2013 and 2013-2024? If so, please clearly state this. Also, in the caption of Figure 5, it sounds like the values are not density-corrected? This needs to be clarified.
- Figure 5: The diverging colorbar is not appropriate for the strictly positive quantity SMB (cf. Crameri et al., 2020, Nat. Commun., The misuse of colour in science communication). Please use a sequential, perceptually uniform colormap (e.g., viridis) instead. Maybe also explain in the caption, that the missing stakes cause the interpolation patterns.
- Z. 302: Which climatological mean was used to fill gaps, the mean for that specific day-of-year?
- Z. 305 f: Please note that mm w.e. yr^{-1} and kg m^{-2} yr^{-1} are numerically equivalent (1 mm w.e. = 1 kg m^{-2}). Since both units are used interchangeably throughout the manuscript (e.g., here sublimation/deposition in mm w.e. vs. the reported SMB in kg m^{-2}), it would help readability to state this equivalence explicitly once, e.g. in Sect. 2, so readers can follow the arithmetic (e.g., 25.61 − 1.47 = 24.14) without unit-conversion confusion.
- Z. 310: This conversion from mass flux to energy flux is not obvious to me: converting the reported net vapor flux (1.47 mm w.e. yr^{-1}) by multiplying with L_s = 2.834 MJ kg^{-1} yields a long-term mean LE of ~0.13 W m^{-2}, roughly an order of magnitude smaller than the stated 'less than 3 W m^{-2}'. Please specify how the value of ***3*** W m^{-2} was derived.
- Figure 6: Consider adding a thicker/highlighted zero line (0.00 mm w.e.) to make the visual distinction between sublimation and deposition clearer at a glance
- Z. 356: "*Without applying the scaling factor, the subsidence of 1.64 m snow layer is limited to 0.05 m"*: please add for context: 'compared to observed/simulated subsidence on the order of tens of centimeters over the same 2008–2024 period', to make the magnitude of the scaling factors' effect explicit or specify the period.

Technical Corrections

- Inconsistent use of "in situ" vs. "in-situ" throughout the manuscript, please standardize to one spelling and follow the journal guidelines.
- Z.91: "in specific period" should be "in a given period" (missing article)
- Z. 213: dashed line (instead of dotted)
- Z. 320: Add the plural "s" to "stake observations"
- Z. 331: 'Kameda' is named as the sole actor, but two references with different first authors are cited together (Takahashi and Kameda, 2007; Kameda et al., 2008). Maybe consider rephrasing.