This manuscript uses three transient ice-sheet-model simulations (PISM-CI, PISM-CESM, and PSUISM) spanning the last two million years to quantify how surface elevation and ice-divide/dome migration at major Antarctic ice-core sites could bias ice-core-based climate reconstructions. The study finds substantial glacial–interglacial elevation variability at East Antarctic sites, a lag between simulated surface-height changes and Antarctic climate signals, and ice-divide/dome displacements of up to ~120 km, with implications for both existing ice cores and prospective "oldest ice" drilling sites. The topic is timely, and the three-model comparison is a genuine strength, but the manuscript would benefit from a number of clarifications, additional justifications, and presentation improvements detailed below. I recommend that the authors address the following comments before the manuscript is considered ready for publication.

<u>Comments on the Abstract</u>

1. "Ongoing prospective future drilling locations" is vague for an abstract headline; naming LDC, Dome Fuji, and Dome A explicitly (as the Introduction does) would make the abstract more self-contained.

<u>Comments on the Introduction</u>

2. The intro lists prior ISM-based corrections (Ritz et al. 2001, Huybrechts et al. 2007, Parrenin et al. 2007b, Cuffey et al. 2016, Sutter et al. 2019) but the gap statement ("earlier ISM simulations... are often on coarse resolution grids and regional and 1D modeling approaches might not capture the full dynamics of the AIS...") could be improved. Consider being more specific: no prior study has compared multiple independent ISMs (and forcing approaches) against each other over the full 2 Myr window at ice-core sites. This narrower claim better motivates the three-simulation, two-model design used here.

3. The three research questions are good, and question (ii) is an interesting one to raise, but it might benefit from a sentence of physical motivation earlier in the text — e.g. why a lag between temperature and elevation would be expected (ice sheet response time). As is, readers unfamiliar with ice dynamics may not see why this matters until the phase analysis section.

4. The individual mechanisms at plateau vs. grounding-line sites are both explained well, but the fact that they produce opposite signs of bias (attenuation vs. amplification) is currently left for the reader to infer. It might strengthen the motivation if this sign-flip were stated explicitly.

5. The phrase "and beyond (Barbante, 2025)" would be clearer with a short clause spelling out what it refers to, since not all readers will immediately recognize the citation. Something like: "...up to 800 thousand years (EPICA community members, 2004) and now extending to roughly 1.2 million years with the

Beyond EPICA ice core (Barbante, 2025)..." would make the claim self-contained without requiring the reader to look up the reference.

Comments on the Methodology

6. PISM-CI vs PISM-CESM justification. You explain PISM-CI's limitation (linear interpolation between snapshots, can't capture spatial pattern shifts) but it would help to state directly why this specific limitation is relevant to elevation/divide reconstructions.

7. Note that PISM (9.8 K/km, dry-adiabatic) and PSUISM (5°C/km, an empirical/environmental lapse rate) use physically different lapse-rate assumptions. The choice of lapse-rate physics differs, which could itself be a major driver of inter-model spread in surface elevation change. Worth flagging this explicitly as a structural difference between the model setups, since it is currently only apparent by comparing separate subsections.

8. Bedrock deformation approaches differ too (viscoelastic in PISM vs. ELRA with a single 3000-yr relaxation time in PSUISM) — again, worth an explicit sentence acknowledging this as a source of divergence.

9. Results throughout are reported as single numbers (m, km, °C, %) without uncertainty bounds — understandable given each simulation is a single deterministic run rather than an ensemble. A brief sentence early in the Methods or Results, clarifying that the inter-model spread reflects structural/forcing differences rather than a formal statistical uncertainty, would help readers calibrate confidence in the reported figures.

Comments on the Results

10. **Section 3.1** — phase-shift estimates (Fig. 4/7) — The magnitude of the reported phase/timing shifts (e.g., "up to 45 kyrs" for PSUISM at the Last Interglacial) is presented as a point estimate derived from a single, fixed bandpass filter (7–250 kyr). It would strengthen confidence in the quoted values if the authors briefly indicated the sensitivity of the phase-shift magnitudes to reasonable variations in the filter band (e.g., a narrower or wider cutoff) or justified why 7–250 kyr was chosen as the appropriate band.

11. **Section 3.1** — The text states "PISM-CI seems to be most in sync with the δ18O record, while PSUISM exhibits the largest phase shift" — but PISM-CI is forced using a climate index that interpolates between climate snapshots "along a paleoclimatic index" (Sutter et al., 2019; Wirths et al., 2025), described as "representative for the climate evolution of the last 2 million years." The Methods do not state what this index is derived from. If it is itself based on an isotope stack (δD/δ18O-type record), then PISM-CI's apparent synchrony with

δ18O may partly reflect the forcing design rather than an independent physical result. A brief clarification of the index's basis, and an explicit caveat if it is isotope-derived, would help readers weigh this specific comparison appropriately.

12. **Section 3.1** — Only one lag value is quantified in the text ("up to 45 kyrs for the Last Interglacial peak in PSUISM"); the rest are qualitative ("shifted by several millennia," "delayed"). Figures 4, 7, and A1 are graphical phase-angle composites, not numeric summaries, so lag magnitudes must be read off plots by eye. A short summary table (lag in degrees and/or kyr, per site, per model, covering both the EDC δ18O and LR04-based analyses) would make Section 3.1's desynchronization claim much easier to evaluate.

13. **Section 3.2** — PSUISM's absence from the ice-divide/dome analysis is noted twice as a scope limitation (in 3.2 section and again in 4.3 section), but neither mention addresses that PSUISM diverges most strongly from the PISM runs in elevation history elsewhere (3.1, 3.3). A brief added sentence — in 3.2 or 4.3 — noting that the reported divide/dome variability estimates (up to ~120 km) are therefore based on two of the three ISMs only, and may not bound the true spread that a PSUISM-inclusive analysis would show.

14. **Section 3.2** — It might be worth considering a summary table here. This subsection reports many individual numbers in prose (~120 km, ~45 km, ~70 km, ~50 km, ~65 km, ~70 km, ~40 km, etc., across six sites and two models), and Figures 5 and 8 are spatial maps rather than numeric summaries. Since Section 3.1 already includes Table 1 for ∆h, an analogous table for ice divide/dome displacement (per site, per model) could improve readability, make it easier for readers to compare sites at a glance, and incidentally make PSUISM's absence from this analysis more visually apparent.

15. **Section 3.2** — The divide-detection method uses a fixed velocity threshold ($\bar{u}max$ = 0.5 m/yr), but no justification is given for this value. It would be helpful if the authors could briefly explain why 0.5 m/yr was chosen or ideally include a short sensitivity check showing how much the reported displacement estimates would change under a different reasonable threshold.

16. **Section 3.3** — Comment (line 240): Grammar/typo: "we observe the same pattern in ∆h variability throughout the the last 800 ka at LDC than at EDC" — double "the," and "than" should be "as" ("...throughout the last 800 ka at LDC as at EDC").

Comments on the Discussion

17. **Section 4.1** — The Discussion states ∆h "has not differed more than 152 m between all three simulations". It's unclear from the text whether this 152 m figure is (a) the maximum inter-model spread at a single site/time point, (b) the maximum such spread across all sites and the full 800 kyr record, or (c) something derived from single-model amplitude ranges in Table 1 (several of

which — e.g. TALDICE/PISM-CESM at ~151 m — are close to this number by coincidence or not). Since this figure is later used to justify the ~10% glacial-interglacial temperature uncertainty at EDC, it would help if the authors stated explicitly, in or near Table 1, how "differed between simulations" is computed and at which site(s) the 152 m maximum occurs.

18. **Section 4.1** — Figure A2 compares the study's simulated EDC Δh against the widely-used Parrenin et al. (2007b) 1D reconstruction. This strikes me as a particularly valuable comparison, since it's the only place in the paper where the results are checked against an independent, previously published, community-adopted reconstruction. Given its importance, the authors might consider moving this figure (or a version of it) into the main text.

19. **Section 4.2** — the structural-difference explanation is convincing but could be tied more directly back to the Dome Fuji divergence noted in 3.3 (opposite-signed long-term trends between PISM-CI and PSUISM). Currently the paragraph explains the general causes (SMB scheme, lapse rate, bedrock deformation, marine ice-sheet physics) and illustrates them with LDC's 150–200 m spread but doesn't explicitly revisit the DF sign-flip as a second worked example.

20. **Section 4.2** — The growth-rate thresholds (13.1 m/kyr, 2.5 m/kyr) are converted into temperature-rate thresholds using a single fixed lapse rate (9.8 K/km), whereas a comparable calculation in 4.1 (WDC deglaciation bias) explicitly uses a lapse-rate range (7–9.8 K/km) to account for uncertainty. Applying the same range here — even though the "secondary importance" conclusion would likely survive it, since the resulting thresholds would still be small relative to observed ice-core temperature trends — would make the argument more methodologically consistent with 4.1 and more robust to the reader.

<u>**Comments on the Conclusion**</u>

21. The Conclusion effectively summarizes the paper's general findings but stops short of a practical takeaway for readers. Since the paper's stated goal is to "assist ice core scientists in the interpretation of their results," it would strengthen the closing to add one sentence ranking which sites/time periods are most vs. least affected by elevation/divide bias — e.g., Talos Dome (3.1: the only site where inter-model Δh differences are not even consistent in sign), EDML (3.2: largest magnitude of dome-position-driven temperature and SMB bias, up to 4°C and >20%), and LDC pre-MPT (4.3: substantially reduced dome variability and bias). This would give readers a quick practical summary without requiring them to reconstruct it from Table 1 and Section 3.