Review of “Nudging an inversion-based ice sheet model towards observed mass change during model relaxation: application to the C.H. Ostenfeld Glacier” by van den Akker et al.

The manuscript presents a nudging approach for the relaxation phase that follows a snapshot inversion in an ice sheet model. During relaxation, basal friction under grounded ice and a basal melt perturbation under floating ice are adjusted based on the misfit between modelled and observed ice thickness and on the modelled thickness change rate. In addition, a correction term based on observed thickness change rates is added to the mass transport equation, so that forward simulations start from the observed thickness changes. The method is implemented in ISSM and applied to C.H. Ostenfeld Glacier in northern Greenland.

The novelty is the extension of nudging, previously applied mainly in spin-up methods, to inversion-based initialization. The approach is conceptually straightforward and addresses a well-known problem: model drift and the inability of initialized models to reproduce observed present-day mass change. The results show that nudging combined with the dH/dt correction keeps the ice thickness and elevation changes close to observations.

The manuscript addresses a relevant problem for the ice sheet modelling community, and I consider it worth publishing. In its current form, however, it would benefit from a discussion of the physical plausibility of the nudged fields, consistent input data, a more rigorous evaluation of the results, and a clear separation between what is demonstrated and what is claimed, as well as from careful proofreading. Most of my comments below concern the presentation and evaluation of the results rather than the method itself, and I hope they help to bring out the strengths of this work more clearly.

General comments

1. Compensating errors and physical plausibility of the nudged fields.

The nudged friction and melt perturbation fields are likely to absorb a range of model and data errors, such as SMB bias, bed errors, the outdated ice-shelf geometry and the fixed front. The nudged fields C_l and P, and their deviation from the inversion result, are not shown. Without them, it is difficult to judge whether the nudged fields are physically meaningful or mainly compensate for other errors. The discussion (L 389-397) itself provides an example of such compensation: to keep the grounding line near its observed position despite the additional mass loss through the outdated ice shelf, the nudging produces an ice stream that is both too slow and too thick. In this case, the grounding-line position is partly reproduced through compensation, with the geometry error being transferred into the friction field. This makes it all the more important to show the nudged fields and discuss their physical plausibility.

2. Inconsistent input datasets and ice-shelf geometry.

The input datasets refer to different periods: the SMB is the 2000-2014 mean, the dH/dt product of Smith et al. (2020) covers roughly 2003-2019, the ice shelf in BedMachine v5 represents the situation around 2007, and no years are given for the MEaSUREs velocity mosaics or the SeaRISE surface

temperatures. For Eq. (1.8) to be consistent, the SMB should match the period of the dH/dt observations. Since the RACMO2.3p2 product used here extends to 2024 (L 94), this should be easy to adjust. I also wonder why a multi-year mean is used. A mean is reasonable during the relaxation period, but the forward simulations could be forced with yearly SMB, or the authors could at least justify the choice of a constant mean forcing. Furthermore, the ice shelf included in BedMachine has largely disappeared today (L 392-393), and the authors themselves identify this as a likely driver of the velocity errors. Would it be possible to adjust the geometry and place the front at its present-day position? Otherwise, this remains a potentially avoidable source of error that may bias the evaluation of the method. The authors may argue that the unmodified BedMachine geometry is commonly used in ice sheet projections; however, since the manuscript explicitly identifies the outdated ice shelf as a source of model error, I think it should be addressed, or at least its effect quantified.

3. Support for the main claims

Some claims in the abstract and conclusions go beyond what is shown by the results. The method is described as a "simpler and cheaper alternative" to transient calibration (L 24-25), but neither a cost comparison nor a direct comparison with a transient calibration approach, such as that of Badgeley et al. (2025), is provided. Likewise, the statements that the method is not model-specific and will contribute to reducing uncertainty in sea-level rise projections (L 25-27) rest on a single glacier with low mass change, simulated with a single model. Moreover, no projections are carried out to demonstrate or estimate that the method actually reduces uncertainty in sea-level rise projections, as was done by Badgeley et al. (2025). In addition, the predictive skill of the initialized model is not yet assessed, for example through a hindcast or by withholding part of the dH/dt or velocity time series from the nudging and using it for independent validation. Finally, the discussion itself attributes the fast convergence to the simplicity of the chosen glacier (L 344-345); this should also be acknowledged as a limitation in the abstract and conclusions. See also my comments on L 67 and L 102/103.

4. Presentation of the main results (Sect. 4.4)

This section contains the key results but is currently difficult to follow; it would benefit from restructuring. (i) The one-time-step results, which demonstrate one of the three main objectives of the method (consistency with observed thinning rates, L 22), are only shown in the supplement (Fig. S2). I suggest moving this figure to the main text, maybe combined with Fig. 6 and with the observations (Fig. 1c) as a reference, using the same colour scale. (ii) The three simulations are described in varying order and with inconsistent names (see also comment on L 306 vs. L 309); a fixed order and unique run names throughout would help. (iii) The forcing of the forward runs (L 313-314) should be described in the methods section. (iv) The results are presented mostly qualitatively ("small", "close to", "largely follows"); quantitative metrics would make the conclusions more convincing.

5. Presentation and readability

The manuscript is difficult to read in its current form because of a number of errors and inconsistencies. These include a sign error (I guess) in the central Eq. (1.8) (see comment on L 168), the use of the same symbol C for both the friction coefficient and the correction term, inconsistent notation for the nudging parameters, parameter values that differ between text and figure legends (H_0 = 100 vs 150 m), incorrect figure references (e.g. L 249, L 300 and L 308), ambiguous names for different simulations (e.g. L 306 vs. L 309), a missing Sect. 4.3, a placeholder in the financial support section (L 445), etc. In

addition, “mass change rates”, “thickness change rates”, “mass-trend change rates” and dh/dt are used interchangeably (e.g. Fig. 1c and its caption), although all figures show thickness or elevation change rates in m/yr rather than mass changes. Most of these issues are minor and easy to fix, but taken together they make it hard to follow the method and to verify the results. I therefore recommend careful proofreading of the text, notation, figures and references. Specific instances are listed in the specific comments and technical corrections below.

Specific comments

- **L 22 / L 67 / L 102-103:** I suggest adding a brief motivation for the choice of C.H. Ostenfeld Glacier, as there are several more prominent glaciers, such as Jakobshavn Isbræ, 79°N Glacier or Petermann Glacier. This would also be a good place to mention the limitations of this glacier system discussed later (L 344-345).
- **L 23 / L 46 / L 61 / L 134-136:** The term “nudging” is used throughout the manuscript but never defined. In data assimilation, nudging usually refers to adding a relaxation term to the prognostic equations that pulls the model state towards observations, proportional to the misfit and with a prescribed timescale. Here, however, the nudging acts on uncertain parameters (basal friction and melt perturbation) rather than on the ice thickness itself, which then adjusts through the model dynamics (as in Pollard and DeConto, 2012). I suggest adding a short definition, ideally in the introduction, that states this distinction explicitly and clarifies how the correction term in Eq. (1.8) relates to it.
- **L 29-30:** That’s not true. According to AR6 Glaciers contribute more to sea-level rise than the GrIS (see Tab. 9.5 in AR6)
- **L 32:** Consider replacing “atmospheric warming” with “glacier surface melt due to atmospheric warming”.
- **L 33-35:** The individual SSP projections are not essential here, since the focus of this work is on model initialization rather than sea-level projections. However, the large spread within each scenario (e.g. 27-354 mm for SSP5-8.5) could be used to motivate the study: part of this spread is attributed to differences in model initialization (e.g. Goelzer et al., 2018, Aschwanden et al. 2021). I suggest either shortening this sentence or explicitly linking the spread to initialization, which would also provide a smoother transition to the next paragraph.
- **L 40:** The phrasing “Ice sheet models typically fail to capture those rates” sounds too harsh, since some models can reproduce historical mass-loss rates. A softer alternative could be: “Capturing historical mass losses remains challenging for ice-sheet models, with implications for sea-level projections…”
- **L 59–60:** Please clarify the goal of the method and define “drift”. As I understand it, the aim is to match the observed thickness changes rather than to remove drift.
- **L 82-85:** The choice of m = 1/3 is justified only by reference to previous studies (Åkesson et al., 2022; Rückamp et al., 2026), the latter dealing with Alpine valley glaciers rather than a marine-terminating outlet glacier. Please provide a physical justification (e.g. m = 1/n following Weertman, 1957). More importantly, Åkesson et al. (2021) show that projections of the neighbouring Petermann Glacier depend strongly on the friction law. Since the friction exponent controls how sensitive the velocity is to changes in C, it likely also affects the behaviour of the nudging scheme and the grounding-line evolution in the forward runs.

Related to this, the authors state (L 83) that C(x,y) accounts for effective pressure, among other processes. With a Weertman law and C kept fixed after nudging, the effective pressure is effectively frozen into C. Near the grounding line, however, effective pressure decreases towards zero as the ice approaches floatation, which should reduce basal friction. In the forward runs with observed thinning rates (Fig. 6a), where the ice thins by up to ~100 m near the grounding line, an effective-pressure-dependent law (e.g. Budd-type or regularised Coulomb) would likely lead to lower friction and faster retreat. The current results may therefore underestimate grounding-line retreat and mass loss. Moreover, during relaxation the nudged C field may absorb changes that would physically be expressed through effective pressure, which complicates its interpretation. I suggest discussing these implications, and ideally testing the method with an alternative friction law.

- **L 86:** I suggest dropping "assumed".
- **L 89-90:** Please briefly justify, ideally with a reference, that these melt rates are typical for northern Greenland.
- **L 93-95:** Since only the 2000-2014 mean of RACMO2.3p2 is used, the details of the underlying reanalyses (ERA-40, ERA-Interim, ERA5) are not needed here. I suggest simply referring to the statistically downscaled RACMO2.3p2 product (Noël et al., 2019) and stating the period it covers (1958-2024).
- **L 95:** Please clarify whether the 2000-2014 mean is applied as an annual mean or as a monthly climatology that retains the seasonal cycle. Given the half-monthly time step, a seasonal cycle would introduce seasonal thickness variations that the nudging scheme could respond to, which may affect the nudged fields.
- **L 97-101:** I suggest dropping "since there is no agreed best approach to model calving rates", as the absence of agreement on calving laws does not rule out including one. The actual reason for the fixed front follows in the next sentence, namely that the nudging scheme targets basal friction and basal melt and cannot adjust calving. I would therefore reorder the paragraph to e.g.: "Our proposed nudging scheme targets basal friction under grounded ice and basal melt under floating ice, and cannot adjust calving rates based on misfits in ice thickness, velocity or thickness change rate. We therefore do not apply a calving law and keep the ice front from advancing. The front can still retreat where ice thins below the 10 m threshold for dynamically active ice."
- **L 102-105:** I suggest moving this paragraph to Sect. 3 (Model domain).
- **L 111:** In Eq. (1.4), u_m and u_o are scalars, whereas here they are referred to as vectors. Please clarify. Out of curiosity, why does Eq. (1.4) only contain the logarithmic term and not also a linear one? As far as I understand from Morlighem et al. (2010), the logarithmic term mainly constrains the slow-moving areas.
- **L 115:** The values of gamma_1 and gamma_2 and the outcome of the L-curve analysis are missing; they could be given in the supplement.
- **L 118-121:** Order of inversion and thermal steady state: if the temperature field is computed only after the inversion, the rheology and hence the velocity change. Which rheology was used for the inversion, and is the ice rigidity B inverted for the ice shelf? What time period is the SeaRISE surface temperature? Why not use the 2000-2014 RACMO product, as for the SMB (or 2013-2019 to be consistent with the observed height changes; also SMB)? Together, these choices may explain why the modelled velocities are too low (Fig. S1).

I also think that the thermal model needs some more explanation (possibly in the supplement). Cold ice or enthalpy model? If enthalpy, is the dynamic BC activated? Which parameters for cold-temperate transition? A figure of the basal temperature and/or the depth-averaged rigidity B would also be helpful.

- **L 138:** “adapted from van den Akker (2026b)”. Can you briefly explain why you dropped the factor “2” in the second term and completely the fourth term compared to Eq. 5a in van den Akker 2026b?
- **L 144-147 and Eq. 1.6:** So, the parameter that is optimized at every timestep, is C(x,y)? If so, shouldn’t it be C(x,y,**t**)? I also recommend clearly differentiating between the friction coefficient C in the initial snapshot inversion and in the nudging procedure (maybe C_0 and C_n, respectively). I was initially confused about which field is optimized, as C(x,y) is defined by the snapshot inversion.
- **L 154:** The value of m_m,0 in Eq. (1.7) is missing in the text. Please use consistent subscripts for m and H, either “m,0” or “0,m”.
- **L 160-163:** Smith et al. (2020) primarily provide surface elevation change rates, from which mass changes are derived using a firn model. Please clarify which quantity is used for C in Eq. (1.8). If elevation change rates are used directly as thickness change rates, firn compaction is neglected, and over floating ice the elevation change must be converted to a thickness change using the floatation condition. The effect of these assumptions could be discussed.
- **L 162:** C(x,y) is the friction coefficient in Eqs. (1.2), (1.5) and (1.6), but C is also the correction term in Eq. (1.8). Since both act simultaneously, this can be confusing, and the correction term should get its own symbol, ideally one containing the subscript “obs” to highlight the link to the observations.
- **L 168:** I think there is a sign error here. In equilibrium, Eq. (1.8) gives −div(bar(u)H)+B=C. Once C is removed, this yields dH/dt=**+C**, not −C. Line 163 defines C as the observed thickness change rate, and L 272 confirms that C is negative for thinning.
- **L 172:** It is unclear whether H_0,m, tau_m and r_m take the same values as H_0, tau and r in the ensemble. The figure captions of Fig. 2 refer to Eq. (1.7) for tau and r, not to Eq. (1.5).
- **L 174:** tau=500 and 5000 years are far longer than the 350-year run time. This choice should be explained. See also L 238-240, where runs with tau = 5000 are practically insensitive to r. Why not test a value of, e.g. tau=10 or 100 years, which are lower than the simulation run time?
- **L 194:** Please indicate Harder and Brikkerne glaciers in Fig. 1.
- **L 196:** The grounding line is not visible in Fig. 1; please add it.
- **L 197-198:** thickening and thinning rates: which dataset? Which period? "Thickening rates" is used for values down to -1 m/yr; thinning is meant. In addition, "near the grounding line" appears twice in the sentence.
- **Figure 1c:** I suggest a colour bar ranging from -1 to 0.1 m/yr with white at zero. Please also label the quantity consistently (see general comment on terminology).
- **Figure 2 (and Fig. 3):** The panel titles “with/no ΔH/Δt correction” are confusing. In Eq. (1.8), ΔH/Δt denotes the modelled thickness change rate, and Eqs. (1.5) and (1.7) also contain a thickness change term, so the titles could be read as if this term were removed from the nudging equations. As I understand it, the panels show runs with and without the correction term C (= observed dH/dt) in Eq. (1.8). I suggest labelling them accordingly, e.g. “with/without dH_obs/dt correction”. Furthermore, since Eq. (1.8) is the continuous mass conservation

equation, ΔH/Δt should be written as dH/dt. Finally, the caption of both figures states H_0 = 100 for the green lines, whereas the legend and the text refer to H_0=150.

- **L 222:** Instead of “most 17m” I would provide the range so that also the low RMSE is reflected, i.e. range from ~8m to 17m.
- **L 228-231:** These lines discuss runs without the correction but refer to Fig. 2a (with the correction). The conclusion that the "correction does not fully prevent long-term drift" then does not follow from the preceding sentence.
  In addition, in Fig. 2b, all nudged runs end at roughly 19-26 m RMSE after 350 years. This is only slightly better than the baseline (~25 m), and the curves are still rising. H75_T500_R0 even exceeds the baseline. The abstract (L 23) and the conclusions attribute the reduction in drift to the nudging. The results rather suggest that the effect only appears in combination with the dH/dt term. This should be stated clearly and discussed.
- **L 249**: The sentence refers to Fig. 2a, but Fig. 3a is meant. The sentence "from 23 to 27 … compared to H_0=500" is hard to follow.
- **L 265:** H150_T500_R0.5 is selected as having “one of the lowest” thickness RMSEs, although the H75 runs in Fig. 2a are clearly lower. The selection would ideally be based on an objective criterion that combines thickness, velocity and thickness change rate misfits, given the trade-off between thickness and velocity errors (L 247-249). Alternatively, or in addition, I suggest running several ensemble members forward, e.g. one for each value of H_0 (the dominant control on the misfits), or the members with the lowest thickness RMSE, the lowest velocity RMSE and the best compromise. The spread among these forward runs, evaluated with the metrics suggested for Sect. 4.4 (General comment 4), would show how sensitive the projected evolution is to the choice of nudging parameters, and whether a low thickness or a low velocity misfit is more important. Such a sensitivity is directly relevant to the claim that the method reduces uncertainty in sea-level projections (see General comment 3).
- **L 282-285 (& L 256):** I think these statements are somewhat too optimistic. Fig. 3a shows that the velocity RMSE rises from ~10-13 to ~23-27 m/yr in all runs, i.e. it roughly doubles. The relaxation itself is costly, even if nudged and baseline runs end up similar.
  In addition, a basin-wide RMSE is dominated by the slow-moving interior. The insets in Fig. 5 show errors of up to about -400 m/yr in the main trunk, where observed velocities are around 900 m/yr. An RMSE for the fast-flowing area (e.g. u > 200 m/yr) or along the main flowline would be more informative.
- **L 300-303:** Neither Fig. 1c nor Fig. S2 shows mass changes; both show thickness (or elevation) change rates in m/yr. Likewise, the statement that “the integrated mass loss rate might be negative and might even be close to the observed value” is not supported by a figure or number. Since the modelled thickness change rates are available, please compute the integrated mass change rate (in Gt/yr) for all simulations and compare it with the observed -0.17 Gt/yr (L 199). In addition, I suggest presenting the thickness change in Fig. 6 also as a mean rate (m/yr) or using the same colour scale as Fig. 1c, to facilitate a direct comparison with the observations.
- **L 300 and 308:** Both refer to Fig. S2c. For the run with dH/dt, Fig. S2a is presumably meant.
- **L 308-311:** The statement that the modelled thickness change rates closely match the observations, and that the nudging “substantially improves the transient mass-balance behaviour of the model”, is based on a visual comparison only. Please support it quantitatively, e.g. by giving the RMSE and spatial correlation between modelled and observed thickness

change rates for all three simulations, as well as the integrated mass change rate compared with the observed -0.17 Gt/yr.

- **Fig. 6:** A mass loss or SLE figure for the 100-year runs would be informative, with caveats, even if the authors do not want to make a statement on committed sea-level rise. In addition, I am confused by the titles stating these are rates (i.e. unit m/yr), while the colour bar title and figure captions state it is the difference (i.e. unit m). Please clarify.
- **L 306 vs L 309:** two different simulations are referred to as “nudging”, which is confusing. I suggest introducing three different simulation acronyms that describe “non-nudging”, “nudging with C”, “nudging without C”. There are other occurrences in the text where it is not clear.
- **L 297-314:** The relation between the single-time-step simulations (L 297-298) and the 100-year forward simulations (L 312-314) is unclear. If the former simply represent the first time step of the latter, i.e. the instantaneous thickness change rate once C is removed, please state this explicitly and present both as one experiment. In that case, the forcing should be described once, before the single-time-step results, or simply refer to Sect. 2 (L 178-181), where the forward simulations are already introduced. Finally, L 297 refers to a single configuration (H150_T500_R0.500), whereas Fig. S2 shows three simulations; please make the wording consistent.
- **L 317-318:** Please change “mass-trend change rates” -> thickness difference
- **L 319:** mass change pattern -> thickness pattern
- **L 330:** Maybe add “ .. as used here as initialization for the nudging approach …”
- **L 333:** "almost near-universal" -> either “almost universal” or “near-universal”
- **L 333-338:** Long sentence. Please split.
- **L 415:** The statement that the study presents “a solution to explicitly control model drift” seems too strong. Without the correction term, the thickness RMSE continues to increase until the end of the relaxation period (Fig. 2b), and several runs with the correction also show a remaining upward trend (Fig. 2a). L 420 itself states that the relaxation “largely” removes model drift. Please rephrase this statement accordingly.
- **L 422:** The forward simulations are described here as “unforced”, whereas L 313-314 states that they are run under constant oceanic and atmospheric forcing. Please correct this, e.g. “a forward simulation under constant present-day forcing”.
- **L 422-424:** This main conclusion is based on a single selected parameter setting (H150_T500_R0.5). Please state this limitation explicitly, or support the conclusion with forward runs from several ensemble members (see comment on L 265). Furthermore, the conclusion that the simulated thickness change “follows the present-day trend” is not sufficiently supported. Fig. 6a shows the cumulative thickness change after 100 years (in m), whereas Fig. 1c shows the observed rate (in m/yr), with different colour scales, so only the spatial pattern can be compared, and only qualitatively. Whether the modelled thinning matches the observed trend in magnitude, and whether it persists, weakens or accelerates over time, cannot be inferred from the end state alone. Moreover, since the forward run starts from the observed rates by construction and is run under constant forcing, a similar spatial pattern is to be expected, at least initially. Please rephrase this statement in line with L 318-319, or support it with time series of the modelled thickness change rates and their misfit to the observations (see comment on Sect. 4.4).

- **L 424-426:** Since the abstract states that the method is not model-specific and would reduce uncertainty in sea-level projections if implemented across models (L 25-27), I suggest adding the application in other ice sheet models to the outlook, in addition to the application to other glaciers. Furthermore, the method currently relies on time-invariant observational fields. A natural extension would be to continue the nudging over a historical period using time-varying observations (e.g. annual velocity mosaics and thickness change time series) and annual SMB forcing, or to withhold such data for an independent validation (see General comment 3).

Technical corrections

- **L 26:** implemenetd -> implemented
- **L 29:** contribitions -> contributions
- **L 88:** caculated -> calculated
- **L 90:** m_deep is not typeset as a formula.
- **L 102:** tributarty **->** tributary
- **L 121:** are ran -> are run
- **L 136:** data assimulation -> data assimilation
- **L 162:** A comma is missing after "velocity vector".
- **L 183:** asses -> assess
- **L 222:** the simulation that have -> the simulations that have
- **L 279:** obsverved -> observed
- **L 292:** obsverved -> observed
- **L 297: Section numbering:** Section 4.2 is followed directly by 4.4; there is no Section 4.3.
- **L 299:** instantanous -> instantaneous
- **L 315:** double occurrence of “in”
- **L 351:** Typo? an tangled -> “a tangled mix” or “an entangled mix”
- **L 415:** explicitely -> explicitly
- **L 440:** All simulation were -> All simulations were
- **L 445:** The placeholder "We need to add the THAWING grant." should be removed.
- **References:**
    - Barnett et al. 2026a and 2026b are the same paper (TC version and preprint). Only the TC version should be cited.
    - Williams et al. (2026) and Jager et al. (2026) lack a journal or source.
    - Inconsistent spelling: "Macayeal"/"MacAyeal", "Deconto"/"DeConto", "Imbie Team"/”the IMBIE team”, "jager"/”Jager”, “van den Akker”/”Van den Akker”.
    - Several references are missing DOIs.