Review of *"Projecting the evolution of the Northern Patagonian Icefield until the year 2200"* by Schaefer et al.

This manuscript presents simulations of the Northern Patagonian Icefield and its future evolution through the year 2200. The manuscript is generally well written, clearly structured, and suitable for publication in *The Cryosphere*. However, I found that the Results and Discussion sections could be strengthened, particularly regarding the treatment and discussion of uncertainties. My main comments are outlined below.

General comments:
==

SMB

If I understand correctly, you apply a present-day surface mass balance (SMB) from Schaefer et al. (2013) and adjust it homogeneously using a temperature anomaly. If so, precipitation changes derived from the Earth System Models (ESMs) are not accounted for. I would expect that under warmer conditions, increased melt would occur but potentially also increased precipitation at higher elevations. Is it not feasible to run the reference experiments using precipitation fields directly from the ESMs you employ?

Related to this, I suggest addressing the following points in the Discussion section:

- Do you expect the SMB forcing approach used here to respond similarly across the range of ESMs considered?

- Are elevation feedbacks included in your SMB–temperature relationship? From the current description, it does not appear that they are.

- Spatial patterns present in the ESMs may not be captured with the adopted methodology. I understand that downscaling ESM outputs may be challenging, but acknowledging this limitation would strengthen the discussion.

- What is the temperature spread of each ESM over the Northern Patagonian Icefield? As I understand it, the standard deviation you apply corresponds to the variability of the *mean* temperature values across models. However, I would expect substantial spatial variability in temperature within the domain. In that case, applying a spatially uniform standard deviation may not fully capture the uncertainty. Would it be more appropriate to use a spatially varying measure of variability instead?

Spin-up

I was also uncertain about certain aspects of the spin-up procedure. As I understand it, you use the SMB field from Schaefer et al. (2013) but increase the input SMB by reducing global temperature by 1°C—is this correct? Additionally, it seems that only one spin-up is performed during 500 years, and then followed by 20-year runs with varying parameters for calibration. In that case, would part of the diagnosed drift between 2000–2020 arise from parameter changes rather than the applied forcing? Some clarification would be useful.

Basal friction

Regarding the basal friction law, you define a relationship dependent on basal temperature and basal water thickness, and Figure 2 shows the effect of varying these quantities. However, no sensitivity experiments are presented for these parameters. The spin-up varies only C_b^0, while C_w, γ, and H_w are fixed. Under this configuration, the relevance of Figure 2 is unclear, and removing it may improve the manuscript's focus.

That said, I believe a sensitivity analysis of H_w or geothermal heat flux would be more informative than focusing solely on friction coefficients, because these could significantly influence basal sliding. Could you include at least one or two sensitivity tests to illustrate this?

Parameter values

I also recommend citing references supporting your chosen values for C_w, γ, and H_w. Similarly, the source of the geothermal heat-flux value should be given. Different choices for geothermal heat flux could warm or cool the bed and potentially alter the dynamical state, so discussing this in the manuscript would be valuable.

Calving law

Concerning calving: How is the calving front represented in the model? Do you employ a level-set method, or do you apply a basal melting rate to ice-front nodes? The applied calving rate of 1000 m/yr appears high. When you say this value "was found to match the current observed calving flux," do you refer to flux magnitude or to the present-day terminus position? If it refers to flux, please provide the corresponding observed calving flux value.

Future scenarios

It appears that only one calibration ensemble member (cal6) is used for the forcing. Relying on a single member may limit the robustness of the conclusions. Since all ensemble members show broadly similar present-day tendencies, could additional members be included to assess the sensitivity of future projections? You could add a weight to these simulations based on their present-day performance. This may also help evaluate the influence of parameters such as friction coefficients or enhancement factors.

Comments on Figures

- **Figures 4 and 5:** Instead of plotting model results alongside observations, it may be more informative to show the anomaly (model − observation). This could highlight spatial biases that are otherwise difficult to identify.

- **Figure 7:** Consider including the simulated ice-covered area for completeness.

- **Figure 10:** It appears similar to Figure 7 but expressed in percentages. If so, you may consider merging or simplifying.

Minor Comments

- **Table 1:** Please clarify the role of the “residual stress parameter.” I could not find further explanation in the manuscript.

- **Line 272:** It should be SSP5-8.5, not SSP1-8.5 (also in Table A1).
- I agree with the other reviewer that removing the ECHAM5 scenario would improve clarity and facilitate the reading of the manuscript.