Reviewer 2

The main idea seems to be that: (a) Geoengineering reduces the effect of global warming, somewhat, on the Iceland ice cap, and (b) this is different from how Greenland behaves for a variety of reasons: thinner, smaller, fewer marine terminating glaciers, high precipitation, high geothermal heat flow, etc. HOWEVER... the main ideas are repeatedly lost in a mountain of detail. Needs major rewrite to put the main ideas (whatever they might be) front and center, and use the details to support it.
Figures need major work. They are unnecessarily hard to interpret. Difficult / inappropriate color scales, lack of titles on axes, overall plot and color scale where they could have been used, Hard-to-distinguish lines with fine gradations in colors, large matching problems between lines and legends, lack of glacier labels (except in one figure), etc. This is making the reader work WAY too hard to understand the data in these figures.
The procedure seems pretty straightforward: run an ESM, use it to force PISM, see what happens. And before that, spinup for a certain period of time. Again, this section needs to be better organized around main ideas, rather than getting lost in the details.

Thanks, we have improved our manuscript significantly according to your valued suggestions. We have added a new section (2.3 SMB modelling) to explain the statistically downscaling method more detailed, and describe how we modelled SMB, and detailed information about PISM parameterizations we used, and a table (Table 1, see reply 76-77) about PISM boundary forcing. These all make our method clearer.

2.3: SMB modelling:
"In this study, the SMB fields used to drive PISM are from Yue et al. (2021), and estimated by SEMIC under the historical, G4, RCP4.5 and RCP8.5 scenarios during 1982–2089. SEMIC in turn is driven by downscaled and bias-corrected ESM data including temperatures, windspeeds, pressures, humidities and radiative forcing terms. We use all CMIP5 and GeoMIP ESM that have complete data fields available, namely BNU-ESM, HadGEM2-ES, MIROC-ESM, and MIROC-ESM-CHEM (Table 1). We statistically downscaled the ESM forcing based on the ERA5 reanalysis dataset (Hersbach et al., 2020). The point of the bias correction is to ensure the mean state of the model parameters matches observations. The separate model trends within each ESM over the observational period remain the same. Thus, the bias correction ensures that models begin close to an observed state, but can then diverge as the separate model climate dictate. The spatial resolution of ERA5 is about 30 km, but still cannot capture the VIC topography. To address this, we first downscaled ERA5 climate to 0.025°×0.025° grid based on their correlation with VIC surface elevation. We find surface elevation is well correlated with near-surface temperature ($R=0.83$, $p<0.01$), downward longwave ($R=0.77$, $p<0.01$) and shortwave radiation ($R=0.74$, $p<0.01$) and specific humidity ($R=0.77$, $p<0.01$), with lapse rates of -5.4 °C km^{-1}, -11.9 W m^{-2} km^{-1}, 15.85 W m^{-2} km^{-1} and -0.59 k k^{-1} km^{-1}, respectively. Precipitation and snowfall are downscaled following De Ruyter-de Wildt et al. (2004). The former is downscaled using Kriging interpolation method, with its empirically exponential relationship with observed surface elevation. The latter is assumed equal to precipitation rate when the daily mean air temperature is below 3°C, otherwise no snowfall occurs. Other SEMIC driven fields (surface wind speed, air density, pressure) are simply bilinearly interpolated due to the relatively minor effects on SMB in SEMIC. Then, we use the downscaled 0.025°×0.025° forcing fields as the

observational reference climate to downscale and bias-correct the ESM fields using the ISIMIP approach (Hempel et al., 2013). The ISIMIP is a trend-preserving approach so that the long-term climate trends in models are preserved, while the mean at each grid cell is matched to observations. There are two fundamentally different ways ISIMIP can do the correction: addition and multiplication, and we follow ISIMIP protocol in deciding which method to use for each meteorological field variable (Hempel et al. 2013). The additive approach is used for most fields preserving, e.g. the absolute changes of the monthly temperature; while the multiplicative method is used for preserving the relative changes for precipitation and radiation. Finally, these 0.025°×0.025° fields were used to drive the SEMIC model. We also bias-corrected VIC surface albedo and considered SMB-elevation feedback in all simulations (Yue et al., 2021). Over the whole VIC, modelled SMB over the period 1991–2010 (Fig. 1d, Fig. 2) is well correlated ($R=0.6$, $p<0.05$) with an interpolated map from 60 measurement sites (Björnsson et al., 2013), although the mean is overestimated by 0.61 m yr^{-1}."

We rewrote the Results section, and added the description of the 4 ESM responses to climate scenarios, which made our story more complete. We also did lots of revisions to the Introduction, Discussion and conclusion parts, especially on the relationship between AMOC and Iceland and geoengineering. Please see our new revised manuscript.

We redraw almost all figures according your nice suggestions, we changed colorbar scale to discrete segments, rather than the high resolution essentially continuous ones that we prefer as they are better able to see differences in response without threshold values, but we accept your suggestion. As for glacier labels, we think only the labeles in Figure 1 are enough, since these are the only ones which occur in the text, and we refer to Fig. 1 everytime they are used in the text. e.g. "Differences between simulated state and measured are mainly at the outlet glaciers of Dyngjujökull, Brúarjökull and Síðujökull (location see Fig. 1; Fig. 3)"

Here are our revised figures:

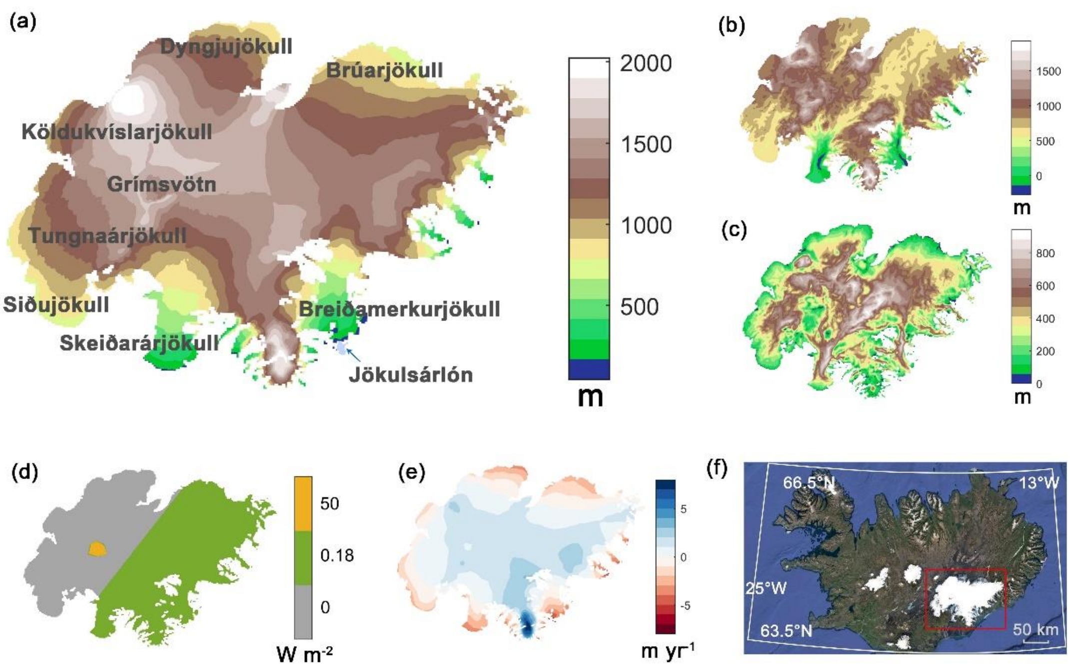

Figure 1. Model input data fields. (a) Vatnajökull ice cap (VIC) surface elevation from Spot5 (data

processing methods see Berthier and Toutin, 2008) in summer 2010; (b) bedrock elevation (Björnsson, 1986; Björnsson and Pálsson.2020); (c) ice thickness; (d) applied upward geothermal heat flux (Flowers et al. 2003), including the Grímsvötn active volcano. (e) annual average surface mass balance 1982-1999 simulated by SEMIC forced by four Earth System Models (Yue et al., 2021). (f) the geographical location of panel (a, red box) observed by Google Earth.

We also showed separate model results, instead of ensemble mean:

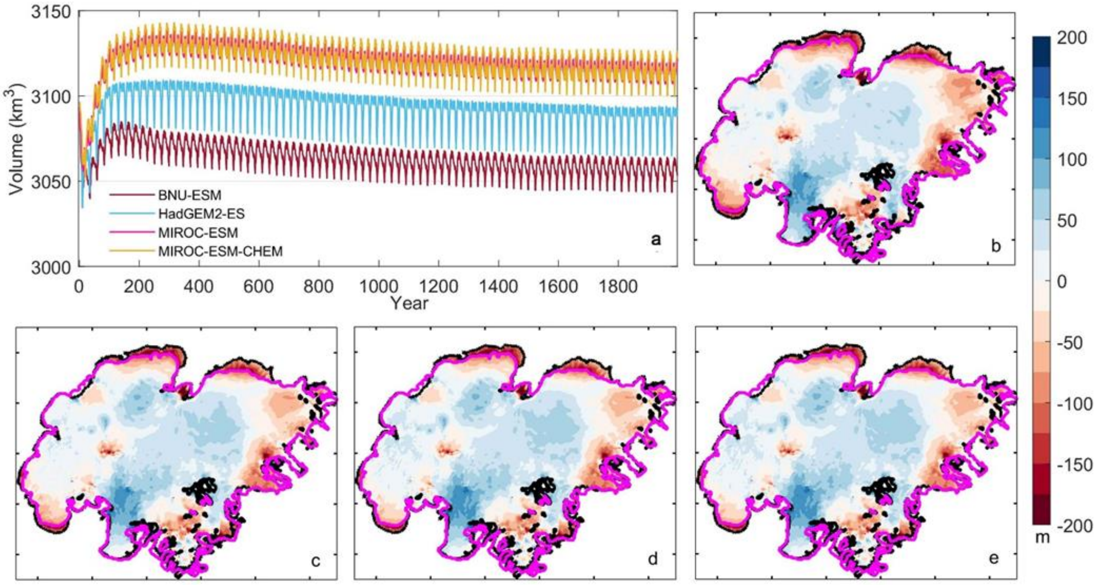

Figure 3. PISM simulated Vatnajökull ice cap (VIC) volume change (a) from the 2000-year climate spin-up simulation driven by repeated monthly SMB fields during 1982–1999 from SEMIC modelling outputs (Yue et al., 2021), driven with downscaled and bias-corrected climate forcings by four Earth System Models. The equilibrium volume is slightly different than present day by -1.3% for BNU-ESM, -0.5% for HadGEM2-ES, and 0.8% for both MIROC models. Subplots (b–e) are the spatial distribution of VIC thickness differences (ice thickness after spin-up minus present ice thickness) from PISM driven by (from b–e) BNU-ESM, HadGEM2-ES, MIROC-ESM and MIROC-ESM-CHEM. The black curves represent the present ice cap extent. The magenta curves represent the extent after spin-up.

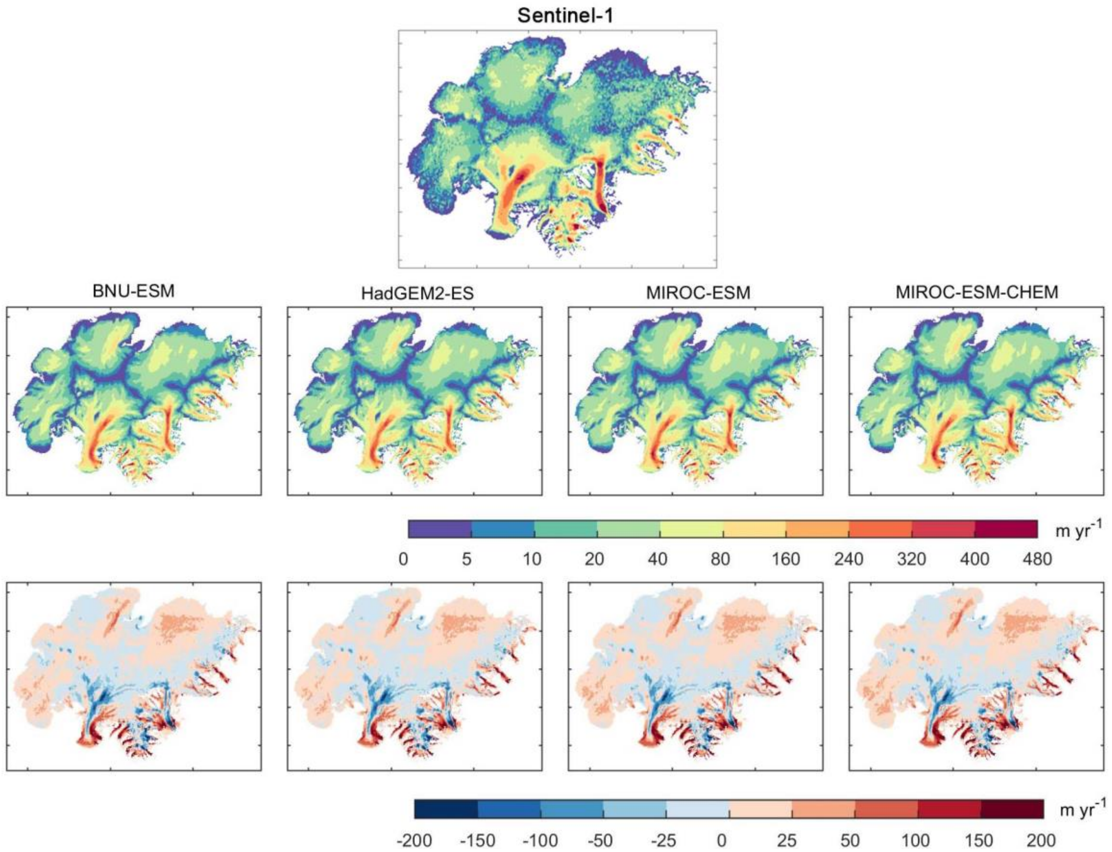

Figure 4. Top: Mean surface velocity over VIC from Sentinel-1, 100 m spatial resolution product (Wuite et al., 2021). Middle row: mean 2015–2020 surface velocities simulated by PISM under the RCP4.5 scenario from the 4 Earth System Model as labeled. Bottom row: the PISM-Sentinel differences.

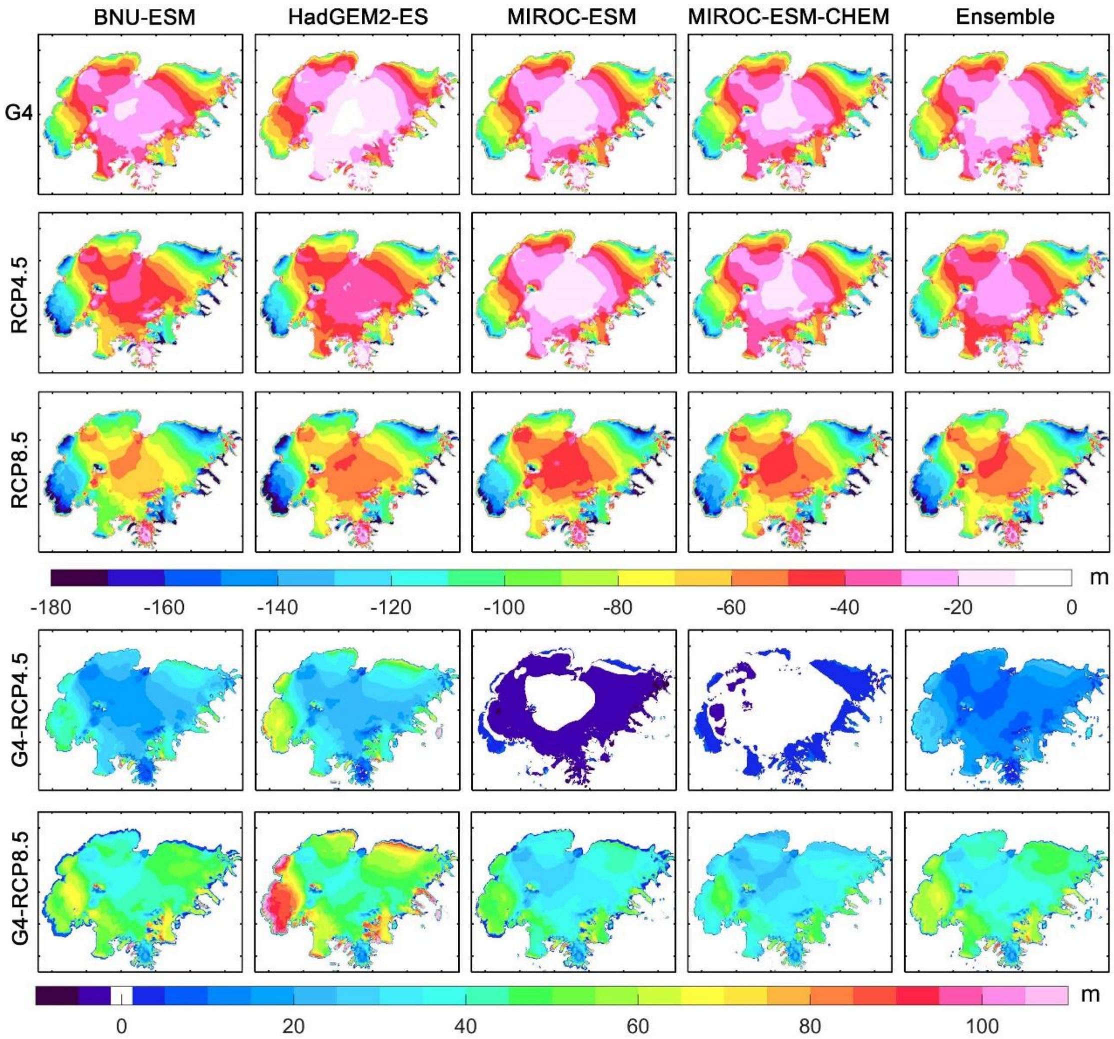

Figure 6. The ice thickness differences from PISM outputs between the year 1982 and 2020 over Vatnajökull ice cap under G4 (1st row), RCP4.5 (2nd row) and RCP8.5 (3rd row) scenarios, and their differences (G4-RCP4.5, 4th row; G4-RCP8.5, 5th row) by Earth System Model (ESM, from left to right), BNU-ESM, HadGEM2-ES, MIROC-ESM, MIROC-ESM-CHEM and ensemble mean. The initial state in 1982 is from each separate ESM.

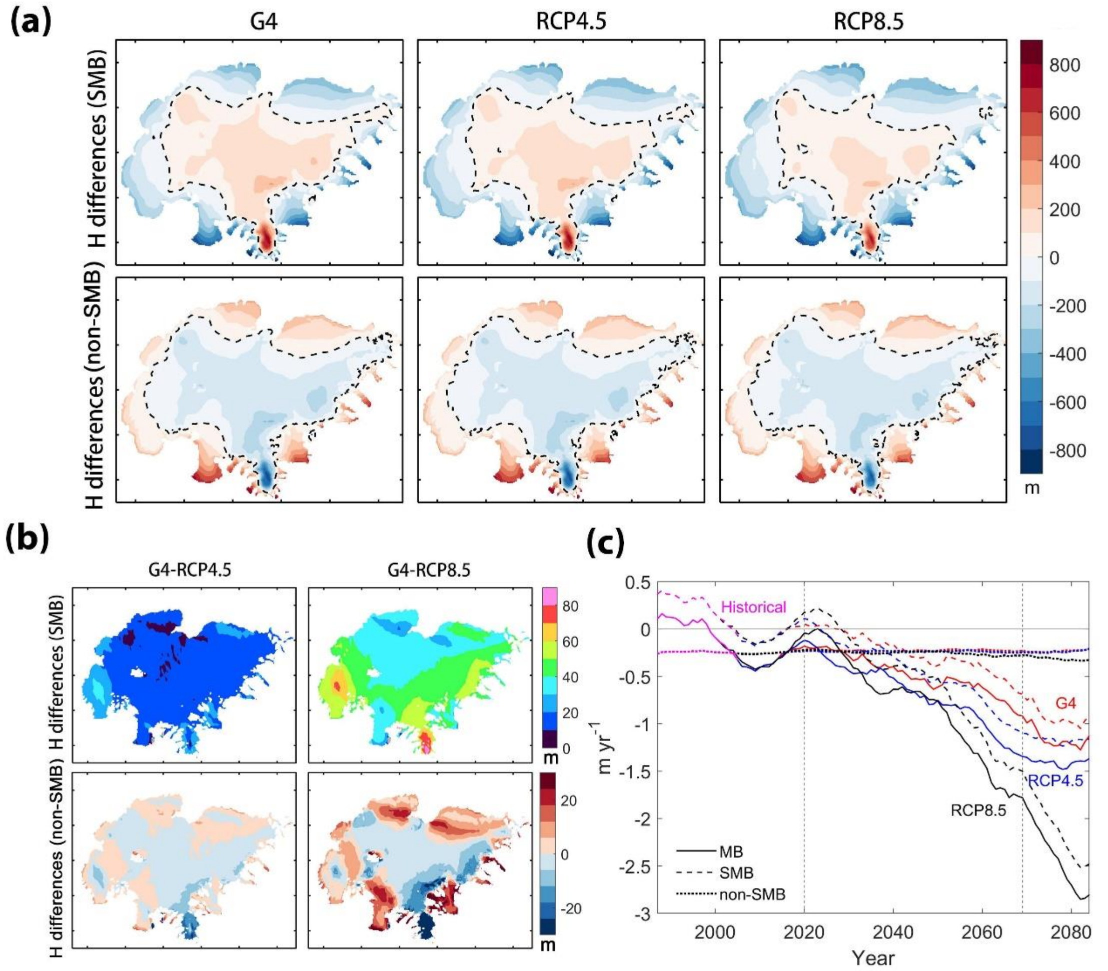

Figure 7. a) Ensemble mean of ice cap height differences, 2089 minus 1982 caused by SMB (1st row) and non-SMB (i.e., ice dynamics and basal melting, (2nd row)) calculated as the difference between MB and SMB, under the G4, RCP4.5 and RCP8.5 scenarios. No change is marked by the dashed black curves. b) Ensemble mean differences (G4-RCP4.5 and G4-RCP8.5) in ice cap thickness by 2089 due to SMB (1st row) and ice dynamics (2nd row). c) Decadal ensemble means of modelled mass balance (solid curves), SMB (dashed curves) and non-SMB (dotted curves) under historical (magenta), G4 (red), RCP4.5 (blue) and RCP8.5 (black) scenarios. The vertical lines denote the beginning and the end of SAI geoengineering. Individual models are in Figs. S2–S5.

38: How do you measure surface MASS balance in meters per year? Is this meters per year of ice? If so, what ice density do you assume? Or is it m/yr of "ice water equivalent?" In which case, why not just stick with mass units and call it 800 kg/m^2? Or is it decrease in elevation with unknown density?

The SMB unit is 'ice water equivalent'. We rewrite the sentence in Section 2.1:

"SEMIC (Surface Energy and Mass balance of Intermediate Complexity; Krapp et al., 2017) is a surface energy and mass balance model that is capable to estimate ice sheet/glacier SMB (unit: water equivalent m yr^{-1})."

51: strike "ever" and "very". Don't use the word "very", ever.

OK. We corrected:

"here we focus only on impacts from SAI on the mass balance of a single ice cap in Iceland. The topic is of wider interest because the behaviour of the North Atlantic under both climate models driven by greenhouse gases, and observational evidence points to a slow-down in AMOC, leading to a much-reduced rate of warming in the North Atlantic relative to the rest of world (Cheng et al., 2013)."

52: How is behavior of North Atlantic atypical? Is it warmer or colder than "typical" and why? "because of the AMOC" is a bit of a mysterious reason.

OK. We corrected:

"Although there are many other potential impacts that might be expected if SAI were undertaken, here we focus only on impacts from SAI on the mass balance of a single ice cap in Iceland. The topic is of wider interest because the behaviour of the North Atlantic under both climate models driven by greenhouse gases, and observational evidence points to a slow-down in AMOC, leading to a much-reduced rate of warming in the North Atlantic relative to the rest of world (Cheng et al., 2013). Under SAI, AMOC slows less than under greenhouse gas climates (Hong et al., 2017; Yue et al., 2021; Xie et al., 2022). Thus, in Iceland, we would expect SAI changes on AMOC and radiative forcing to have compensatory effects to the ice cap. Furthermore, the Arctic warmed 6 times faster than the global mean from 1998-2012 (Huang et al., 2017), leading to concerns on the stability of the Arctic cryosphere, and examination of possible roles for geoengineering methods in its preservation (Lee et al., 2021).Whether SAI might even lead to exacerbated ice mass loss in the North Atlantic is an important question that goes to the fundamental reason for ever doing SAI – that is does SAI better preserve the important elements of the current climate system than plausible greenhouse gas emissions scenarios?"

54: What do words like "considerably" and "much less" mean??

'considerably' means the AMOC slows sharply. 'much less' means the degree of AMOC weakens in future under geoengineering (Hone et al., 2017; Yue et al., 2021) less than without geoengineering. We rewrite:

"Thus, in Iceland, we would expect SAI changes on AMOC and radiative forcing to have compensatory effects to the ice cap. Furthermore, the Arctic warmed 6 times faster than the global mean from 1998-2012 (Huang et al., 2017), leading to concerns on the stability of the Arctic cryosphere, and examination of possible roles for geoengineering methods in its preservation (Lee et al., 2021).Whether SAI might even lead to exacerbated ice mass loss in the North Atlantic is an important question that goes to the fundamental reason for ever doing SAI – that is does SAI better preserve the important elements of the current climate system than plausible greenhouse gas emissions scenarios?"

55: "Furthermore, polar amplification leads to concerns on the stability..."; and add a reference for polar amplification.

We rewrite:

"Furthermore, the Arctic warmed 6 times faster than the global mean from 1998-2012 (Huang et al., 2017)"

57: Langue seems overly wordy

We rewrite:

"Furthermore, the Arctic warmed 6 times faster than the global mean from 1998-2012 (Huang et al., 2017), leading to concerns on the stability of the Arctic cryosphere, and examination of possible roles for geoengineering methods in its preservation (Lee et al., 2021)"

58: Not sure what effects are unwelcome? The descriptions of geoengineering effects earlier in the paragraph seem welcome. I'm confused...

Unwelcome impacts mean the geoengineering may fail in this region due to the Arctic amplification and the impact of enhanced AMOC under geoengineering. We changed to:

"Although there are many other potential impacts that might be expected if SAI were undertaken, here we focus only on impacts from SAI on the mass balance of a single ice cap in Iceland. The topic is of wider interest because the behaviour of the North Atlantic under both climate models driven by greenhouse gases, and observational evidence points to a slow-down in AMOC, leading to a much-reduced rate of warming in the North Atlantic relative to the rest of world (Cheng et al., 2013). Under SAI, AMOC slows less than under greenhouse gas climates (Hong et al., 2017; Yue et al., 2021; Xie et al., 2022). Thus, in Iceland, we would expect SAI changes on AMOC and radiative forcing to have compensatory effects to the ice cap. Furthermore, the Arctic warmed 6 times faster than the global mean from 1998-2012 (Huang et al., 2017), leading to concerns on the stability of the Arctic cryosphere, and examination of possible roles for geoengineering methods in its preservation (Lee et al., 2021).Whether SAI might even lead to exacerbated ice mass loss in the North Atlantic is an important question that goes to the fundamental reason for ever doing SAI – that is does SAI better preserve the important elements of the current climate system than plausible greenhouse gas emissions scenarios?"

58: "Effects" is a fine word, no reason to use "impacts." Impact is a verb when one body hits another and makes a crater... Turning that into a noun and using instead of an existing good word ("effects") is torturing our poor English language.

'impact' can be a noun, and the usage is as popular as verb. See oxford English dictionary: https://www.oed.com/search?searchType=dictionary&q=impact&_searchBtn=Search

And see IPCC report https://www.ipcc.ch/report/ar6/wg2/ "The Working Group II contribution to the IPCC Sixth Assessment Report assesses the impacts of climate change, looking at ecosystems, biodiversity, and human communities at global and regional levels."

60: strike "state-of-the-art." This is a science paper, not marketing material. Instead of "state-of-the-art," just state which version of PISM you used. eg; PISM 3.5 (or whatever).

Ok. We changed:

"We simulate the response of the VIC with the Parallel Ice Sheet Model (PISM; version 1.0) driven by monthly SMB from 1982–2089 under CMIP5 historical (1982–2005), RCP4.5 (2006–2089), RCP8.5 (2006–2089) and SAI G4 (2020–2089) scenarios."

63: Just say **"business as usual" scenario**. I understand it's a scenario we won't like. But it's an oxymoron to label "business-as-usual" as "extreme", especially if it's the most likely scenario. In other words, your use of the word "extreme" is poorly defined here.

No, we disagree. The emissions commitments given by the countries that have signed up to the Paris 2015 accord produce a trajectory closest to the RCP4.5 scenario (Kitous and Keramidas, 2015). Countries in general follow international agreements and commitments because the alternative is chaos. Hence RCP4.5 is the most likely scenario. It is precisely because countries have recognized that "business as usual" will be extremely costly and have decided to cap their emissions. Business as usual is an extreme because it assumes no mitigation, and just about every country accepts the need for mitigation, the question is how much will be done. Thus, business as usual bounds one end of the emissions trajectories, i.e. it is an extreme. However, we rewrite this sentence:

"RCP4.5 (Thomson et al., 2011) is a stabilization scenario with emissions similar to those agreed under the Paris 2015 agreement (Kitous and Keramidas, 2015), while RCP8.5 (Riahi et al., 2011) is a "business-as-usual" scenario that is a likely outcome if we do not make any efforts to reduce the greenhouse gas emissions."

63: Not sure of the word "branches off? Maybe use "depart" instead? Anyway, why are we using a scenario that started in 2020, when it's now 2022?

Branching off is the standard terminology used in climate modeling for beginning a new scenario from an existing one. E.G. https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019MS002009 and countless other articles. G4 has been run by various climate models. There is a wealth of data available to analyze. This article is, we thought, clearly an academic exercise to examine possible impacts of SAI on ice in the North Atlantic. It is not an operational forecast. There are no significant differences in climate across scenarios in the first few decades of the 21st century, it makes no practical difference that we are in the year 2022 already.

67: Did you really use PISM version 1.0? Please explain which vesrion of PISM you ACTUALLY used. Also, please use correct PISM URL: https://www.pism.io

Yes, we do use PISM 1.0, although the newest version is 2.0.2. We changed the URL as your corrections.

72: Are they "free parameters" or "hyperparematers"? Are the two terms different, and which is more appropriate here?

Sorry we don't understand this question as we have never heard the term "hyperparematers". According to google they exist in relation to machine learning, which has no relevance here. They are parameters in the sense that we understand the meaning of the word parameter. We are using it in the standard way.

73-74: Do you have references for any of these schemes?

We added references and the description for each parameterization.

The PISM model (version 1.0; Bueler and Brown (2009); https://www.pism.io) is an open-source ice sheet thermo-dynamic model that has been used in numerous studies of a wide range of ice sheets and glaciers (e.g., Aschwanden et al., 2019; Yan et al., 2020). The evolution of the ice cap surface elevation

H is calculated by mass continuity equation:

$$\frac{dH}{dt} = M - \nabla \cdot \vec{Q} - M_b, \tag{2}$$

where *t* is the time step, *M* is the mass balance, M_b is the basal melt rate, $\nabla \cdot \vec{Q}$ is the ice flux calculated by stress balance model. PISM provides several parameterizations to describe the ice stress balance, ice flow, basal sliding and ice calving (details see PISM manual; https://www.pism.io/docs/). The choices of parameterizations and free parameters followed Schmidt et al. (2020), and validated the simulations using observations over Vatnajökull. In brief the parameterizations we used in this study are:

—We use hybrid stress balance model (Bueler and Brown, 2009) with both Shallow Ice Approximation (SIA; Hutter, 1984) and Shallow Shelf Approximation (SSA; Morland, 1987) to solve ice vertical deformation and longitudinal stretching, allowing simulation of both slowly flowing ice cap interiors and fast flowing outlet glaciers.

—Ice rheology is parameterized by Glen's flow law (Glen, 1955):

$$\tau = 2\eta D\,, \tag{3}$$

where τ is the deviatoric stress tensor, *D* is the strain rate tensor, and η is given by:

$$\eta = \frac{1}{2} A(T)^{-1/n} d_e^{(1-n)/n}, \tag{4}$$

where the parameter *A* is strongly dependent on ice temperature, d_e is the second invariant of the strain rate tensor, flow exponent *n* is commonly taken the value of 3.

—Ice front calving rate *c* is calculated by Eigen scheme (Levermann et al., 2012):

$$c = K \cdot \max(0, \epsilon_{\parallel}) \cdot \max(0, \epsilon_{\perp})\,, \tag{5}$$

where *K* is a constant that explains the ice properties relevant to calving, $\epsilon_{\parallel}$ and $\epsilon_{\perp}$ denote the strain rate along and transversal to horizontal ice flow, respectively.

—Basal sliding is estimated by a pseudo-plastic law (see PISM manual; https://www.pism.io/docs/), which estimates the basal shear stress τ_b through the yield stress τ_c, basal velocity u_b, and the parameters of velocity threshold $u_{threshold}$ and power *q*:

$$\tau_b = -\tau_c \frac{u_b}{u_{threshold}^q |u_b|^{1-q}}, \tag{6}$$

To initialize PISM over VIC, we need surface elevation, bedrock altitude, upward geothermal flux, ice temperature, and monthly SMB for boundary conditions (Table 1, Fig. 1, Fig. 2). We re-gridded these data to a 500 × 500 m resolution in order to balance the computation costs and a more realistic ice surface velocity. As with Schmidt et al. (2020), we prescribe temperate ice at 0°C for the whole VIC, as Icelandic glaciers are temperate (Björnsson and Pálsson, 2008). Upward geothermal flux is from Flowers et al. (2003) and assigns a value of 0.18 W m^{-2} in the eastern sector of VIC, zero in the western sector, and 50 W m^{-2} for the Grímsvötn active volcano (Björnsson, 1988). Surface elevation is from observations by the Spot5 satellite (Berthier and Toutin, 2008), which represents the period from June to September in 2010. Bedrock topography is based on radio echo profiles since 1980 (Björnsson, 1986, Björnsson and Pálsson, 2020).

76-77: Where / how did you get all these datasets? Don't hide the references in the Figure captions.

Ok. We made a table to be clearer.

Table 1 A summary input data fields in PISM.

PISM input fields	Data source	Period	PISM running resolution	Reference
Surface mass balance	SEMIC output driven by downscaled and bias-corrected climate fields from [a] BNU-ESM, [b] HadGEM2-ES, [c] MIROC-ESM, [d] MIROC-ESM-CHEM	1982–1999, repeated for 2000 years (PISM spin-up) 1982–2005 (CMIP5 historical) 2006–2089 (RCP4.5, RCP8.5) 2020–2089 (GeoMIP G4)	Monthly; 500 ×500 m	Yue et al. (2021)
Surface elevation	Spot5 satellite	June to September 2010	500 ×500 m	Berthier and Toutin. (2008)
Bedrock topography	Radio echo profiles	1980	500 ×500 m	Björnsson, (1986); Björnsson and Pálsson. (2020)
Ice cap thickness	Surface elevation minus bedrock topography	—	500 ×500 m	—
Upward heat flux	Assigns typical values	—	500 ×500 m	Flowers et al. (2003); Björnsson. (1988)
Ice temperature	Prescribed 0 °C everywhere	—	500 ×500 m	Schmidt et al. (2020)

[a] Ji et al. (2014), [b] Collins et al. (2011), [c,d] Watanabe et al. (2011).

83: What is .025deg^2 resolution, in km, for Iceland?

It's about 1.2 km. We added the km resolution: 0.025°×0.025° (about 1.2 km×1.2 km).

Figure 1:

i. (d) and (e) should be reversed; as generally things are laid out from left-to-right, not in circular or spiral patterns.

Actually often things are laid out in clockwise or anti-clockwise directions. But we changed them as wished.

ii. Use standard terminology such as ELA, instead of "equilibrium line boundary."

ok

iii. For (e), why is it so artificial?

We assume you mean why is there a sudden jump in geothermal heat at the volcanic hot spot? The temperature scale is not smooth or linear (see your other comments).

There are no heat flux observations at the Vatnajökull, so we followed settings from Flower et al. (2013), we set 0 (western sector, subsurface hydrothermal circulation is believed to be sufficiently vigorous to prevent heat from reaching the base of the ice), 0.18 W m^{-2} (eastern sector), and 50 W m^{-2} (volcanic hot spot).

iv. Use a better color scale, i.e. (a) one made for elevations, and (b) a scale with discrete segments, not continuous. It's hard to tell what's what with contiguous scales. See here for color schemes suitable for topography: http://soliton.vm.bytemark.co.uk/pub/cpt-city/views/topo.html

http://soliton.vm.bytemark.co.uk/pub/cpt-city/ncl/tn/topo_15lev.png.index.html

ok

iv. In camption, but the Berthier & Toutin reference AFTER "in summer 2010"

This is the correct place since we cite a methods paper for the SPOT5, nothing to do with Iceland. The data for Iceland has not been published, so we only cited the methods. We changed:

"**Figure 1.** Model input data fields. (a) Vatnajökull ice cap (VIC) surface elevation from Spot5 (data processing methods see Berthier and Toutin, 2008) in summer 2010;"

vi. For (d), use a +/- gradient color scale (again, with discrete colors not continuous). Blue/red is good; with blue as SMB>0 and red as SMB<0. This is a good one:

http://soliton.vm.bytemark.co.uk/pub/cpt-city/gery/tn/seismic.png.index.html

http://soliton.vm.bytemark.co.uk/pub/cpt-city/ncl/tn/temp_19lev.png.index.html

With a properly centered color palette, you won't need to draw the ELA as a dotted line.

Thanks for your suggestions. We revised Figure 1:

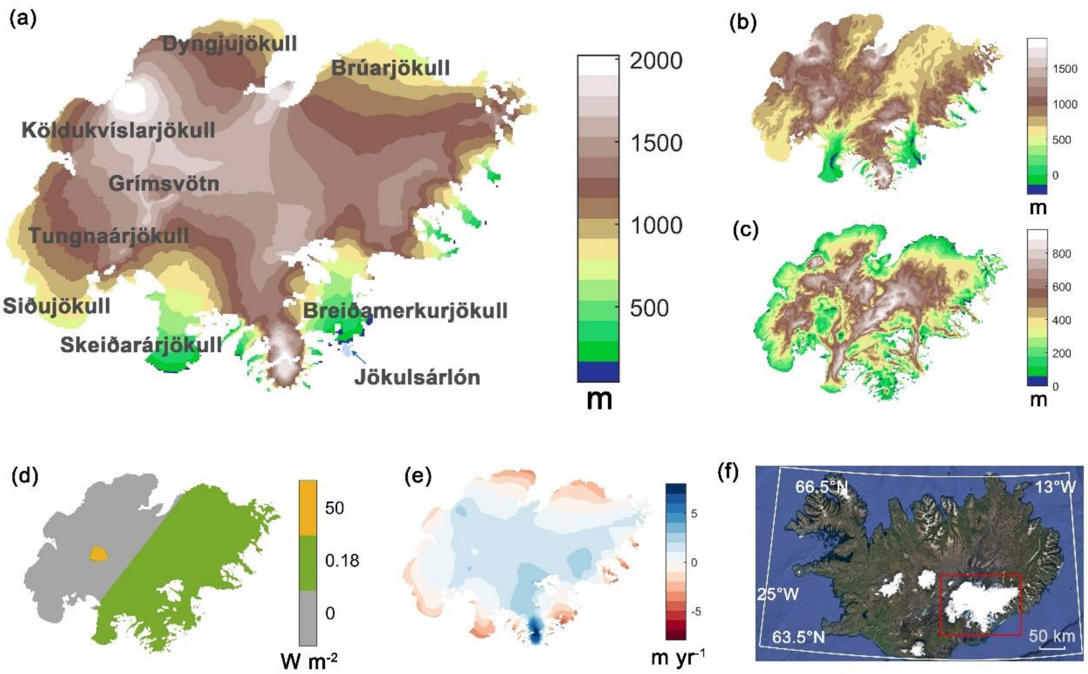

Figure 1. Model input data fields. (a) Vatnajökull ice cap (VIC) surface elevation from Spot5 (data processing methods see Berthier and Toutin, 2008) in summer 2010; (b) bedrock elevation (Björnsson, 1986; Björnsson and Pálsson.2020); (c) ice thickness; (d) applied upward geothermal heat flux (Flowers et al. 2003), including the Grímsvötn active volcano. (e) annual average surface mass balance 1982-1999 simulated by SEMIC forced by four Earth System Models (Yue et al., 2021). (f) the geographical location of panel (a, red box) observed by Google Earth.

93: Why spinup for 200 years? Figure 2 seems to suggest 400 is sufficient. Or why not 4000 years?

Not exactly right. If you carefully look at volume in 400 years and 2000 years, it does show difference!! (about 15 km^3). Actually, when spin-up after 1500-year, volume shows little change, so no need to extend to 4000 years.