The authors are grateful to the referees for the through feedback and constructive comments provided in the review process. All comments have been addressed, as specified in the table below, and have resulted in significant improvements to the manuscript.

Referee#1 Comments	**Author response**
Lines 67-71: Good to add this small paragraph as summary of the novelty. I suggest removing lines 70 and 71, so finishing the paragraph at " […] trusted snow model."	Agree. The text in lines 70-71 was redundant and has been removed.
Line 58: I think it should be snowpack	This has been changed
Line 139 "is A temporal limitation"	This has been corrected.
Line 140: Full stop at the end.	Full stop added
Line 151. "model parameters THAT must be specified"	Correct. "that has been added"
Line 152. I guess it's topography, not topology.	Correct. Topology changed to topography
Line 155. I think it should be snowmelt (Also in line 267 and 269).	Agree. Snow melt changed to snowmelt throughout
Line 156: was followed.	This has not been changed. This sentence refers to the recommendations of Andersson, 2006 and should be in plural.
Line 158: were, not was.	This has been changed.
Line 159: 10x10 meters?	Correct. This has been changed.
Table 1. (and anywhere else if I missed it) Units must be written exponentially as per The Cryosphere submission guidelines.	Correct. Units have been revised to be written exponentially throughout the manuscript.
Line 166: "… number OF times A model grid cell … "	This has been changed.
Figure 2b and 2c: I asked the authors to provide a plot of number of observational stations in time and the mean snow depth of those stations over time. This is now in Figure 1 right panel. It can be seen that before 1960 there were only around 5 stations measuring snow depth over Iceland. This dramatically increases after 1960. Those few stations before 1960 seem to have a shallower snow depth than the stations in the period after 1960 (judging by the dark line in the plot). If a trend is computed over the yearly average of all stations in Iceland, this can give an inaccurate trend. Perhaps those 5 stations were located at coastal Iceland (lower snow depth), and the stations from 1960 are located more inland (higher snow depth). The trends from 1930-2021 would therefore be artificial if it is computed over the average snow depth of all stations for each year. The correct thing would be to do the	Good points. The trends and their estimated significance of the trends illustrated in Figure 2b and 2c presented in Table 2 have been updated based on the historical period 1950-2021 as you are correct that the number of measurements is much lower and less representative prior (mostly costal locations). We have also made the same period changes to Figures 2b and 2c so as to a) show only collectively representative snow data as suggested by the referee and the number of IMO stations reporting snow data significantly increased in the 1950's (as illustrated in Figure 1) and b) the period starting in 1950 corresponds both to the NEX-NDDP data shown in Figure 2a) and the period of the SNOW-17 simulations performed in the study.

average trend, instead of the trend of the average. That means, calculate a trend for each of the IMO stations, then show the average trend. This should be done for the results in Figure 2b and 2c. Alternatively, trends could be computed only from 1960, with higher confidence that the distribution of stations is not changing significantly.	We have made a note of the collective representativeness of data prior to 1950 in the text describing Figure 2 in Lines 202-204
Line 208: Why temperature in K? I think it should be in Celsius.	Agree. This has been changed.
Figure 3 caption: specify which model simulations are these.	Good point. The caption to Figure 3 has been revised to include specifications about model simulations.
Figure 4: Even though the authors have addressed my comment here, something still looks odd. The colour scale of SCF change in 4a goes from +100% to -100% (is this supposed to be per decade?). Judging from the colours, some areas in Iceland are quite close to +100% change over the period 2001-2021. This seems too much since values in table 2 show that the MODIS trends are in the order of 2-4% per decade, which sounds more sensible. I believe either the labels or some calculation is wrong.	Table 2 shows the average SCF trend across the country for the same period which smooths out the variation observed in Figure 4. So, although some pixels may have large changes in SCF over the period the average change across the country will be much lower. Also, the color scale in Figure 4 colors the relative SCF over the whole period 2001-2021 within in each pixel, not the ∆SCF. This has been clarified in the text Lastly, The 100% values on the color scale were chosen to encompass all observations and for simplicity. We note that 100% is uncommon and the highest values observed in Figure 4 are related to areas where recent surface changes are responsible for the change in SCF, e.g. at the termini of the outlet glaciers where the receding glacier has caused SCF to drop considerably from the historical 100%. This has been clarified in the text.
Line 260. I suggest to rephrase: "… , their trend is opposite. The observational data show an increasing SCF trend while the simulations show a decreasing trend over the historical period 1950- 2020."	Good point. Thanks! This sentence has been rephrased
Line 269: "leading to A countrywide"	This has been changed

Referee#2 Comments	**Author response**
L46: I was surprised to see the introduction of the term "regional circulation model" since it is completely new to me and seems to be	The text has been revised to refer to Regional Climate Models

uncommon in the literature. I would thus suggest that the authors switch this to "regional climate model". If you are worried about consistency you can also switch "general circulation model" to the (roughly) synonymous "global climate model".	
L85: The word "parameter" gives the impression of something that is more static than the dynamic properties that together describe the seasonal snowpack, such as snow depth, SWE, SNC, and SNCM. I would thus encourage the authors to instead refer to these as "variables" throughout the text. This would also be in line with the GCOS use of the term essential climate variables (ECVs) among which many snow variables feature.	Good point. The text has been revised to refer to "variables"
Figure 1: I do not recommend representing the number of stations as bar plots which (in my opinion) are rarely a good idea. The same information would be more clearly conveyed as line plots. This would also make the snow depth line plot more visible.	Good point. The plot has been changed from bars to lines.
L101: Be more specific here and change "infrared" to "shortwave infrared" since snow is still quite reflective in the near infrared. Consider also adding a reference to Dozier (1989) who was (to the best of my knowledge) the first to suggest the NDSI (without calling it that).	Good points. Thanks. The text has been revised to refer to "short wave infrared" and the reference to Dozier has been added.
L102 Change "indicates" to "often indicates", otherwise you are overlooking false positives (i.e. NDSI> 0 when FSCA=0) which, despite being less frequent than true positives in cold regions, do occur.	This has been changed.
L108 For the advantages of spectral unmixing with higher resolution sensors I would recommend a reference to Cortés et al. (2014) as well	Good point. The reference has been added
L118: Although it is perhaps implicitly obvious, consider adding that the CMIP6 version could be considered in future studies.	This has been added
L141: Change "is temporal" to "is a temporal". Delete the second sentence here ("It limits. . . ") since there are many other applications of this data that are limited by polar night. If you want to keep it, say "such as" or provide a few more examples.	These changes have been made.
L147: Thanks for clarifying the model resolution. As I understand it you are focusing on modeling the seasonal snowpack, yet nowhere does it state	The glaciers were not masked out in the observations nor the simulations. Figure 4 shows the RGI glacier outlines for reference but the

that you masked out glaciers in the observations (both MODIS and SNCM). I suspect this masking is done somehow, otherwise you would have a representativeness error (see e.g. Janjić et al., 2018) in your observations, but this needs to be stated somewhere in the manuscript.	MODIS data did include the glaciers (SCF = 100). The SNCM data is specifically collected by the IMO for mountaintops observed from mostly lowland areas some of whom do not loose all snow every year. However, as most of the glaciated areas are in the interior highland and not directly visible from most IMO sites this would not cause a significant error.
L151: This suggests that the SNOW-17 algorithm was coded in GEE by Anderson (2006), which is of course not the case. Instead, you could just change this to "The SNOW-17 algorithm (Anderson, 2006). . . "	Good point. This has been changed and the reference to Anderson 2006 moved for clarification.
L155: Change "parameters must" to "parameters that must".	This has been changed
• L159: Change "as followed as it incorporates" to "were followed as they incorporate".	This has been changed
• L162: Clarify what "10 × 10 DEM" means in this context. I guess you mean 10 m resolution?	Correct. This has been clarified.
As a side comment on Table 1, it is laudable that you seek to justify your choice of parameters with references and ancillary datasets. Nonetheless, this exercise is somewhat misleading as these are effective parameters that compensate for the physical processes that are missing in a degree day model. As such, they do not really have universal values that can be extracted from look-up-tables (or similar) based on ancillary data. Instead, their "optimal" values will likely vary considerably based on the forcing data used and other details of the experimental setup such as the spatial resolution of the model. My issue is that this presentation may (inadvertently) camouflage the uncertainty that exists in these parameters rather than embracing it and attempting to calibrate them with the data (e.g. MODIS snow cover) that you have at hand. This becomes especially apparent when one of two references in Table 1 is from a conference abstract, was there no subsequent peer-reviewed publication by these authors that goes into more detail on their methods? In summary, although I understand that calibration is beyond the scope of this work, the uncertainty introduced by the particular choice of parameters should at least be mentioned somewhere in the manuscript.	Good point. This has been clarified in the text in lines 159-162 .

L174: This sentence is problematic. I don't see why published datasets would be less uncertain than derived datasets that seek to make improvements. Please change the formulation here. It is fine to say that for simplicity and to stay within the scope of your study you used these datasets without further manipulation.	Agree. This sentence has been revised. L174.
L180: To me at least, this is not an explanation of what your p-values measure. Instead, you have merely passed the problem along to the term "95% confidence level". Their are plenty of definitions floating around, see e.g. Ambaum (2010) or Benjamin et al. (2018), and I would urge the authors' to adapt the wording in these to their particular tests. The reason I recommend this is that null hypothesis significance testing plays a central part in your results (the term significant is mentioned 4 times in Section 4 alone), so I believe it is important that you properly define what is meant by this term.	Good points. Thanks. This sentence has been revised in line 182-184 for clarification.
L182: Change "ensemble average" to "ensemble mean".	This has been changed.
L270: Please use × not ∗ for multiplication when using scientific notation, i.e. $p = 1.54 \times 10{-}5$.	This has been changed.
L277: Since you are using "significant" in a very specific technical sense elsewhere, I would recommend changing this (different) usage of the word to "considerable" or similar.	This has been changed.
L290: Unsure how helpful it is to introduce the acronym CESM2 which you then never use elsewhere.	Good point. The acronym has been removed.
Table 3 (and elsewhere): I am not sure why you are citing Hall et al. (2006) since you are using the V006 MODIS snow cover product not the V005 product. The correct citation is surely Hall et al. (2016) throughout the manuscript. Also consider following the suggested citation for Hall et al. (2016) on the product webpage at NSIDC from which yours deviates slightly.	Good points. This has been changed.