We thank the reviewer for the comments, which have improved the paper.

General comments

1. Section 4.2:

SEB evaluation in ERA5, ERA-Interim and RACMO2.3: this is a unique feature of the paper, as already commented by another reviewer it is "one of the first validation studies of the ERA5 reanalysis over the Greenland Ice Sheet". Yet only a half a page paragraph is reserved for the presentation/discussion of this evaluation, with all the tables shown in the Supplementary Material. While I do understand that this is not the actual primary scope of the paper, I think that the manuscript and the readers would benefit from a more exhaustive discussion of the evaluation, maybe including a summary figure/table in the main text. Perhaps some of the lengthy descriptions of the surface energy fluxes (Section4.1.2) could be shortened. Similar results and discussions can be already found in previous works (e.g. Kuipers Munneke et al., (2018)) while the evaluation of ERA5, ERA-Interim and RACMO surface energy fluxes is a novelty. Please don't misunderstand me here, I am not suggesting to completely change the scope of the paper or its structure but just maybe to revise the balance of the results section.

Reply: A recent ERA5 evaluation in Greenland was published in Delhasse et al., (2020), and we now make a reference and brief comparison with that study.

We use the results presented in the previous section to evaluate T_{2m}, albedo, radiation fluxes, Q_h and Q_l in ERA5, ERA-Interim, and RACMO2.3p2, the latter forced at the lateral boundaries by ERA-Interim during 2003-2018. We compute model output at the AWS locations using an average distance-weighted interpolation method using the four nearest grid points. Evaluation of KAN_L, KAN_M, KAN_U, THU_L and THU_U are included in the Supplementary Materials, and the evaluation of S5, S6, S9 and S10 can be found in *Noël et al., (2018)*. Tables S2-S5 (In the Supplementary Materials) show the root mean square error (RMSE), the mean bias (MB) and the correlation coefficient (R) based on linear regressions on daily observations of the PROMICE AWS.

Although ERA5 better represents the observations than ERA-interim, the improvement is not statistically significant for all the near-surface variables, in agreement with *Delhasse et al., (2020)*. For Q_h and Q_l, RACMO2.3 provides the highest correlations. For THU_U (Table S3), RACMO2.3 shows high correlation coefficients for shortwave fluxes and 2 m temperature, and Q_h and Q_l are also relatively well represented with correlation coefficients between 0.8 and 0.7, higher than both ERA reanalyses. For albedo, ERA5 outperforms ERA-Interim at most stations. This is likely caused by the new snow albedo scheme, which changes exponentially with snow age in ERA5, and resets fresh snow albedo, while ERA-Interim set a maximum constant albedo for snow events (*ECMWF, 2016*).

We conclude that the regional climate model RACMO2.3 remains a useful addition to reanalysis products for the simulation of GrIS near-surface climate and SEB.

2. Validation of RACMO melt flux:

This is probably the main problem the paper has in its current version. Nowhere in the manuscript I have found a validation of RACMO modeled melt flux, yet this variable is used extensively in the discussion section. In line 449 the reader is referred to Noël et al., (2018),whose SEB evaluation at sites S5, S6, S9, and S10 includes the melt flux as well (e.g. Tables 2-5). Ultimately modeling melt accurately is one of the final goals of any surface energy balance model and regional climate model used to study ice sheets changes. An often proposed and used approach to this type of validation is to evaluate point studies (e.g. like the SEB modeling results here presented) against in-situ observations and then evaluate regional climate models against the point studies. The current manuscript already uses this framework for all the SEB components, but in my opinion it needs to include the same analysis for the melt flux as well.

Reply:

(1) A validation of SEB model against in-situ ablation rate observations is presented in Figure 5 (Average 10-day modeled and observed ice melt). This only works for ice with known density, i.e. at KAN_L, KAN_M, S5, S6 and THU_L which are situated below the equilibrium line. See for the detailed descriptions Section 3.2 (SEB model evaluation) and discussion of Figure 5.

(2) We added a row in the evaluation tables in which RACMO2.3p2 melt rate is compared to that from the SEB model for station KAN_L, KAN_M, KAN_U, THU_L and THU_U. Previously, an extensive evaluation of RACO2.3p2 ablation rate with all available observations from Greenland S5, S6, S9 and S10 was performed by *Noël et al., (2018)* and also showed good agreement.

Specific comments

L52-58: I think there are other studies in the literature worth mentioning that specifically address the surface energy balance on the Greenland Ice Sheet (e.g. Charalampidis et al., (2015), Vandecrux et al., (2018)).

Reply: We added the suggested literature in the "Introduction".

Charalampidis et al., (2015) use a surface energy balance model forced by five years of K-transect AWS measurements to evaluate the seasonal and interannual SEB variability, in particular the exceptionally warm summers of 2010 and 2012. *Vandecrux et al., (2018)* present a simulation of near-surface firn density in the percolation zone, to quantiy the influence of climatic drivers such as snowfall and surface melt.

L76-88: the goals of the paper are not clearly stated before the structure is outlined. This paragraph could be streamlined to make the actual goals more straightforward. E.g.: we study the SEB at two transects ... we put these results into a broader context using these products ... which are validated in this way ...

Reply: We re-organized this paragraph as follows:

We study the dependency of west Greenland SEB and melt on large-scale circulation variability along two GrIS AWS transects, i.e. the southwestern Kangerlussuaq (K-) transect and the northwestern Thule (T-) transect. We put these regional results into a broader spatial context using reanalysis (ERA5, ERA-Interim) products and output of a regional atmospheric climate model (RACMO2.3). ERA5 is the latest reanalysis product from the European Centre for Medium-Range Weather Forecasts (*ECMWF; Dee et al., 2011; Hersbach and Dee, 2016*), and replaces ERA-Interim, considered to be the leading product over GrIS until now (*Albergel et al., 2018; Bromwich et al., 2016*). Because both the PROMICE and IMAU AWS are not assimilated in ERA5, these data can be used to assess its quality and that of regional climate models. Thus, we also include an evaluation of ERA5/RACMO2.3 SEB components over the western GrIS.

Table1: I wonder if here ELA is used instead of elevation, also are some of the weather stations discontinued now? Would it be possible to put the full operational periods? (e.g. Start Date - End Date).

Reply: we changed ELA to elevation. For the data periods of every station used in this study, we put them in Figure 2.

Table 1 AWS location, elevation and start of observations

Station	Latitude(N)	Longitude(W)	Elevation (m a.s.l)	Start Date	End Date
S5	67.08	50.10	490	27/08/2003	01/01/2019
S6	67.07	49.38	1020	01/01/2003	01/01/2019
S9	67.05	48.22	1520	26/08/2003	27/08/2019
S10*	67.00	47.02	1850	17/08/2010	13/09/2016
KAN_L	67.10	49.95	670	01/09/2008	18/02/2018
KAN_M	67.07	48.84	1270	02/09/2008	18/02/2018
KAN_U	67.00	47.03	1840	04/04/2009	19/08/2018
THU_L	76.40	68.27	570	09/08/2010	05/10/2018
THU_U	76.42	68.15	760	09/08/2010	06/09/2018

*S10 is currently stopped while other stations are still operational.

Section 2.2.1: are the AWS data processed? E.g. is any correction applied to the datasets before being used as model input? If yes processing procedures should be described here or referenced appropriately.

Reply: For the data processing, we have cited references from the data providers IMAU (*Smeets et al., 2018*) and GEUS (*Fausto et al., 2012b; Van As et al., 2011*).

Snow and ice height records cannot always be used directly to assess sensor height changes because of AWS design changes and/or settling of the structure. For PROMICE AWS, we use the results from a physically based method to remove air-pressure variability from the signal of the pressure transducer records (*Fausto et al., 2012b; Van As et al., 2011*). For details of S5, S6, S9 and S10 data biases, corrections, and data gap filling in case of sensor failure, we refer to *Smeets et al. (2018)*.

L121: to my understanding "emitted longwave radiation" is not used to drive the model but in the model evaluation.

Reply: Thank you, we removed "...and emitted..." and added at the end of the sentence: "and emitted longwave radiation is used to evaluate the model performance."

L123-124: "where temperature is recalculated to the reference height of 2 m using the SEB model" this procedure should be better explained or an appropriate reference given, if it's important.

Reply: we reformulated as follows:

The height of the temperature/humidity sensor continuously changes due to ablation and/or accumulation and settling of the station. In order to compare to model output at the 2 m reference height, AWS temperature and humidity are recalculated to this height using the flux-profile relations applied to the turbulent fluxes from the SEB model. To illustrate the data time series at the nine AWS, Figure 2 shows the full record of 2 m temperature.

L127-133: what is the surface height measurement used for? For model evaluation, as stated below, but is it also used as precipitation input for the model? It would be good to state what this data is used for and then describe its limitations and corrections applied.

Reply: Information about the surface height is required for turbulent flux calculations, to identify surface type for albedo, to feed snow accumulation into the model and for correction of wind, temperature and humidity to standard heights. We added:

The sonic height ranger provides changes in the surface height, which allows us to accurately determine snow depth, surface type (ice/snow) for albedo, sensor height required for turbulent flux calculations as well as for correction of temperature and humidity values to standard height.

L134-139: the phrasing of this paragraph could be improved to better deliver the message.

Reply: we rephrased as follows:

Note that AWS time series have differing lengths and completeness. For model evaluation with surface temperature (Fig. 4) we used all available hourly values of emitted longwave radiation, i.e. data points used for Figure 4 coincide with the time series as shown in Figure 2. The evaluation using observed ice melt (Fig. 6) uses data starting in 2008, to maximize overlap between the various AWS time series. For the calculation of the average SEB seasonal cycle we used only complete years (Tables 3 and S1, Figures 7, 8, 9 and 10).

Figure 2: is this figure really needed just to show the data availability period?

Reply: We think Figure 2 is necessary, to demonstrate in a single overview how AWS time series have differing lengths and completeness.

Section 2.2.2 and 2.2.3: maybe these two sections could be merged to improve the readability of the paper. Also the title should include the ERA-Interim product.

Reply: The title of section 2.2.2 is changed "2.2.2 ERA-Interim and ERA5" as requested. But we prefer not to merge Sections 2.2.2 and 2.2.3, because the former describes Global reanalysis products whereas the latter describes Regional climate model products.

L185-186: convoluted sentence, readability could be improved.

Reply: The sentence is corrected as requested.

The Surface Energy Balance (SEB) model uses AWS data as input. It iteratively solves for the value of Ts for which the energy budget is closed.

L192: what does the author mean exactly with the surface value of the calculated subsurface heat flux?

Reply: We reformulated as follows:

..., G is the subsurface heat flux, evaluated at the surface, and**...**

L215-217: the process of Smeets and van den Broeke., (2008) could be better explained or simply skipped (e.g. ... following the study of Smeets and van den Broeke., (2008) a value of ...).

Reply: We simplified this paragraph as requested.

Following the study of *Smeets and van den Broeke., (2008)* a z0 value of 1.3 $*10^{-3}$ m is chosen for S5, S6, and KAN_L when ice is at the surface, and 1.3* 10^{-4} m when snow covers the surface at these AWS sites.

L226-233: how is the subsurface part of the model initialized?

Reply: We add more information about the initialization as follows:

The subsurface model is initialized using measured density and temperature profiles at the date of station installation, and assuming no liquid water.

Figure 4: any comment on the fact that it appears that modeled surface temperature is 0C much more often than the observed, this could mean that modeled melt is overestimated.

Reply: The extent to which this holds is hard to say, because both in model as in observations temperatures are capped at the melting point. Considerable uncertainty also exists in the 'observed' surface temperature. Given that observed and modelled ice melt agree well, a systematic bias in calculated melt is not supported.

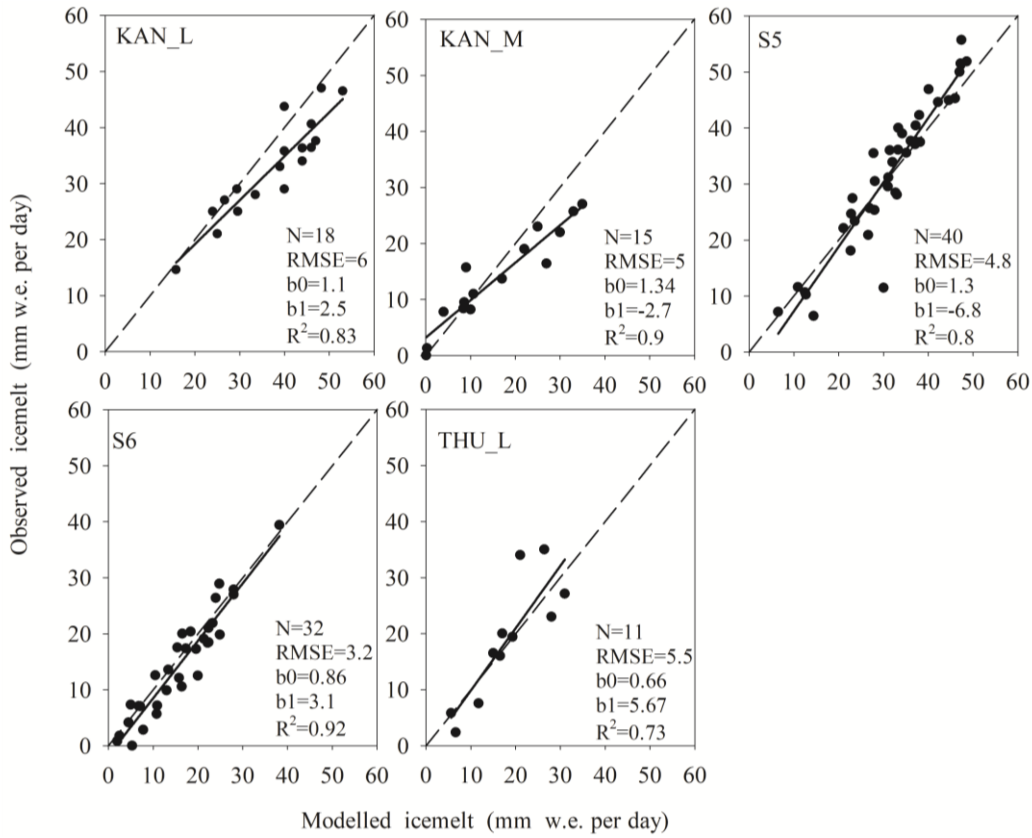
L251-252: I find this sentence a bit misleading, there is still information to be retrieved from 0C surface temperature (e.g. see previous comment about Figure 4), however it is true that the amount of melt cannot be assessed just by using melting surface temperature as a proxy.

Reply: To clarify, we reformulated as follows:

When temperature reaches the melting point, it no longer varies in time and as such it can no longer be used to evaluate SEB model performance. Instead, we assess...

Figure5: modeled data and observed data axis are inverted compared to Figure 4, this should be avoided at all cost. Be consistent with the chosen convention (e.g. modeled data always on the x-axis).

Reply: Here Figure 5 is fixed to keep modeled data always on the x-axis.



New Figure 5

L275-279: such generalization should be avoided in my opinion when site characteristics vary so much. E.g. at S5 the radiation penetration effect on total cumulative melt flux is neglectable but what about at a much higher elevation site like S10 where the melt flux is much smaller?

Reply: Here we add another study about Greenland summit to give a radiation penetration range from lower station to the summit for the model uncertainty analysis.

Van den Broeke et al. (2008b) and Kuipers Munneke et al., (2009) used a spectral albedo model based on the parameterization by *Brandt and Warren (1993)* to calculate subsurface penetration of shortwave radiation at S5 and at Greenland Summit station. Subsurface melt was only found to be important at S5, but with little influence on the total melt. Based on these results, here we assume that that neglecting subsurface radiation penetration in the SEB calculations has little effect on the total cumulative melt flux.

Section 4.1.1 and Figure 6: why there are no model values of surface height change for KANU and S10? Also doesn't the model simulate accumulation?

Reply:

As the SEB model does not simulate accumulation, there are no model values of

surface height change for KANU and S10 which are situated above the equilibrium line.

Why are measured height changes compared to modelled ice melt in Figure 6?

Reply:

As we don't know the density of snow and firn which have melt, the SEB model present the modeled ice melt by assuming an ice density of 910 kg/m^3. That's the reasons of measured height changes compared to modeled ice melt in Figure 6.

Also some of the dashed lines are not continuous (e.g. S6, KANL, ...) why is this the case? A better explanation should be given here. (The nature of this comment is similar to my general comment about RACMO melt flux validation)

Reply:

The dashed lines are not continuous due to the gap data of the model input, as shown in Figure 2.

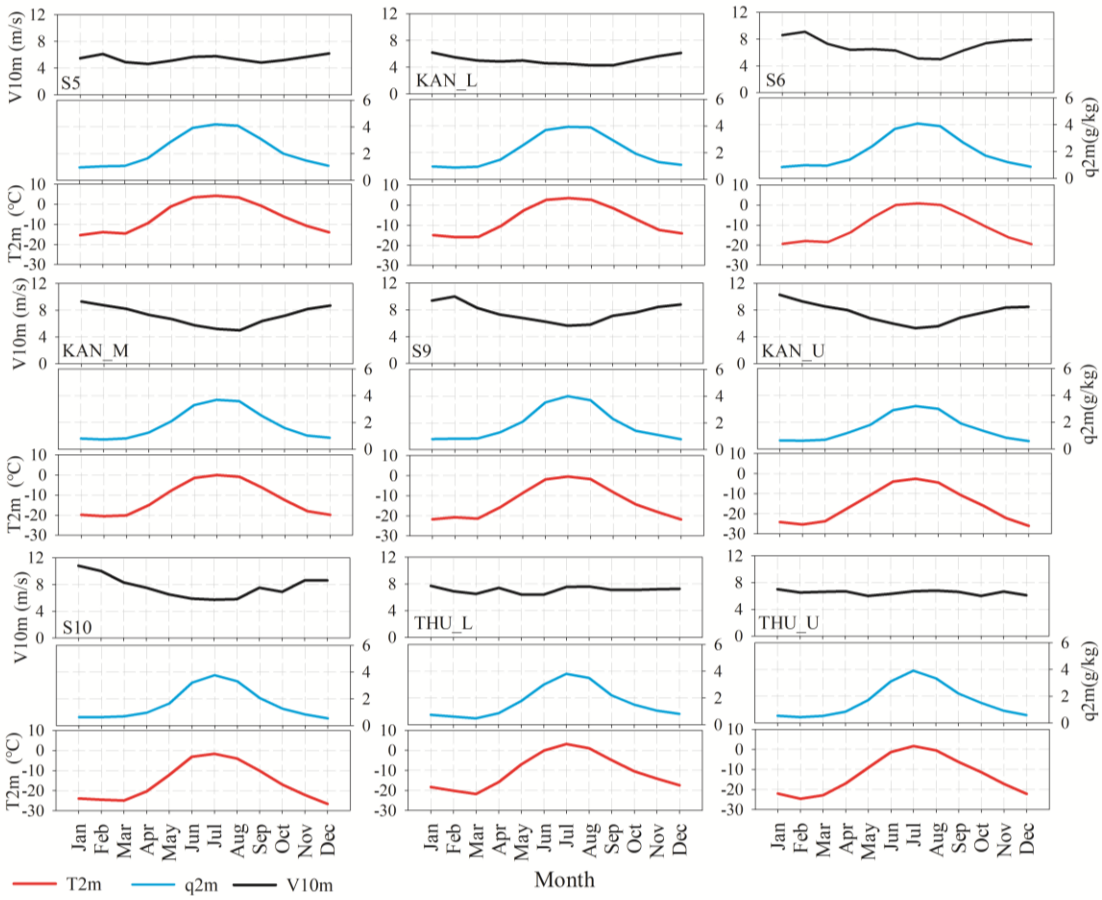
L309-311: reference about the cloud cover product used here?

Reply: Here we use cloud cover data estimated from PROMICE AWS based on L_{in} and air temperature. We added a reference for this equation.

Probably owing to more frequent and thicker clouds along the T-transect (cloud cover 0.51 at KAN_L vs. 0.56 at THU_L in summer, using cloud cover estimates from PROMICE AWS based on Lin and air temperature according to (*Favier et al.*, 2004).

Figure8: I would rather put T2m and q2m on the same subplot (but different axis) since they are correlated rather than q2m with the wind.

Reply: Combining your suggestion with the suggestion from the first reviewer, we decided to separate the plots of the three variables. If q2m and t2m are plotted on the same subplot, it is difficult to distinguish the magnitude of the amplitude in q2m.

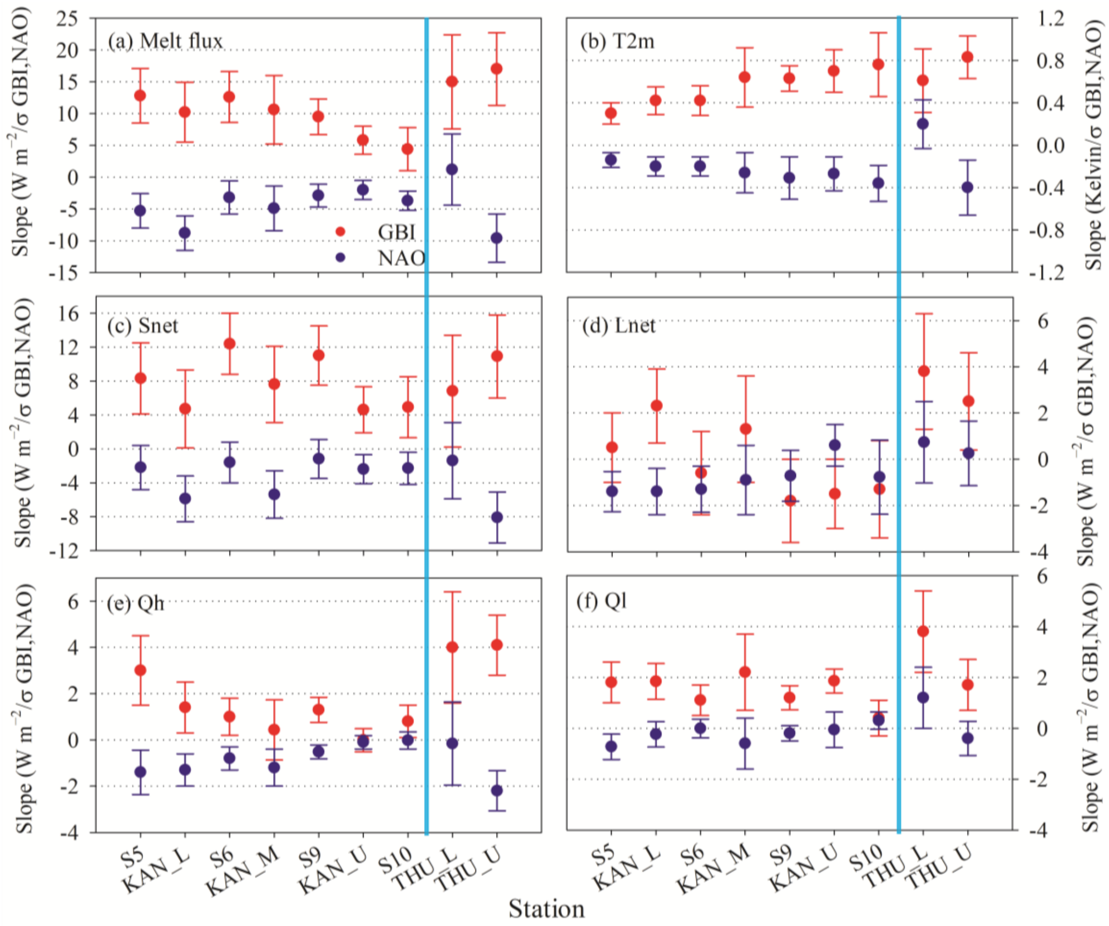


New Figure 8

Figure 11: this figure needs a bit of work from the reader to be fully understood. Providing additional descriptive text (e.g. at L470) to assist the reader would help. Also consider keeping the y-axis symmetric and with the same range. This would help in assessing the difference between different fluxes.

Reply:

We now added the slope unit for every subplot to assist the reader. But if we keep the y-axis range for (c), (d), (e) and (f) same with Fig11 (a) (-15 ~25 W $m^{-2}/\sigma_{GBI,NAO}$**), then it will difficult to distinguish (d), (e) and (f), because the ranges of (d), (e) and (f) are -4~6** W $m^{-2}/\sigma_{GBI,NAO}$. **As a compromise, we decided to make Fig11(d), Fig11(e) and Fig11(f) with the same y-axis range instead.**



New Figure 11

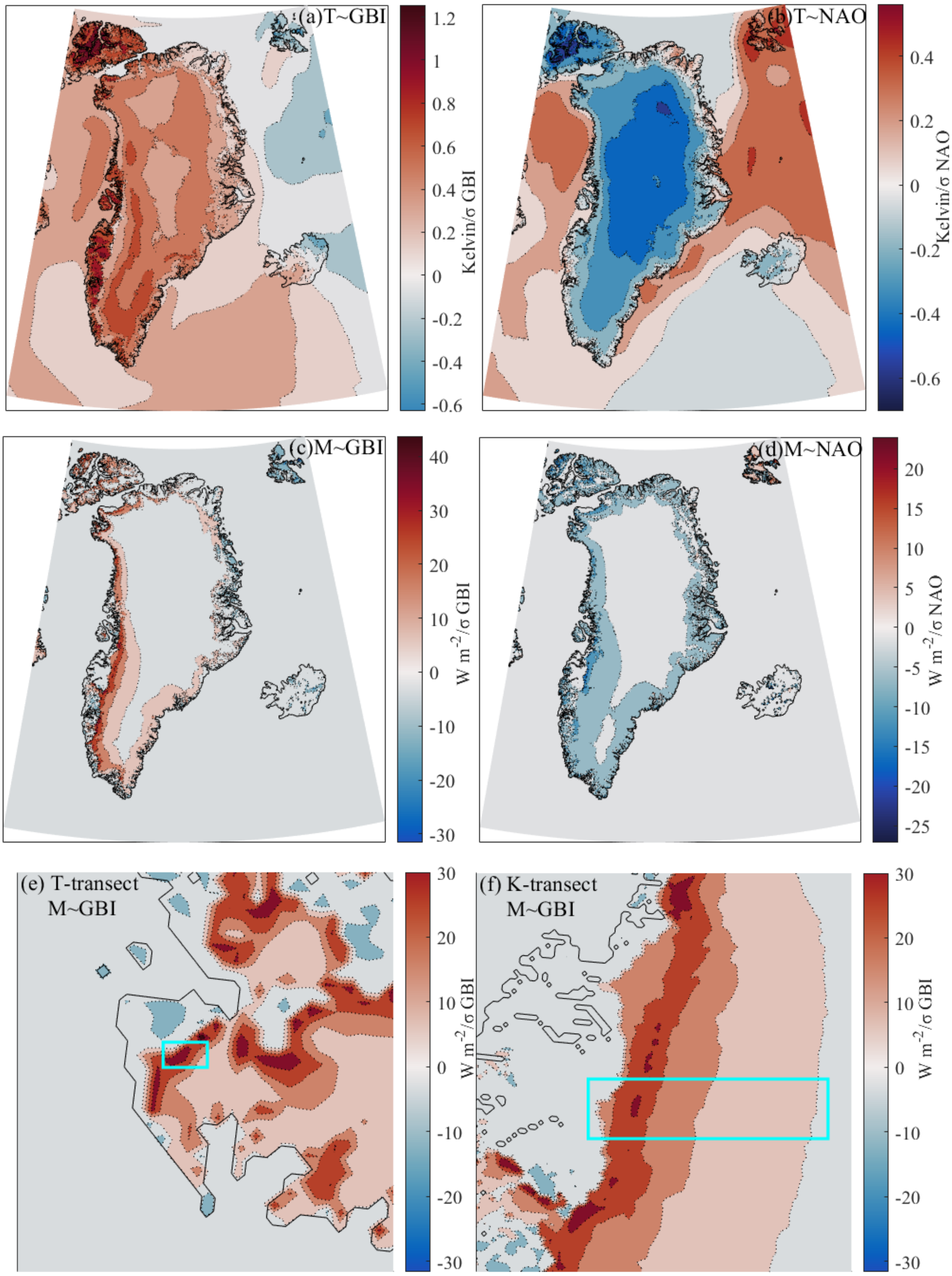
L512-515: convoluted sentence, readability could be improved.

Reply: we changed this sentence as requested.

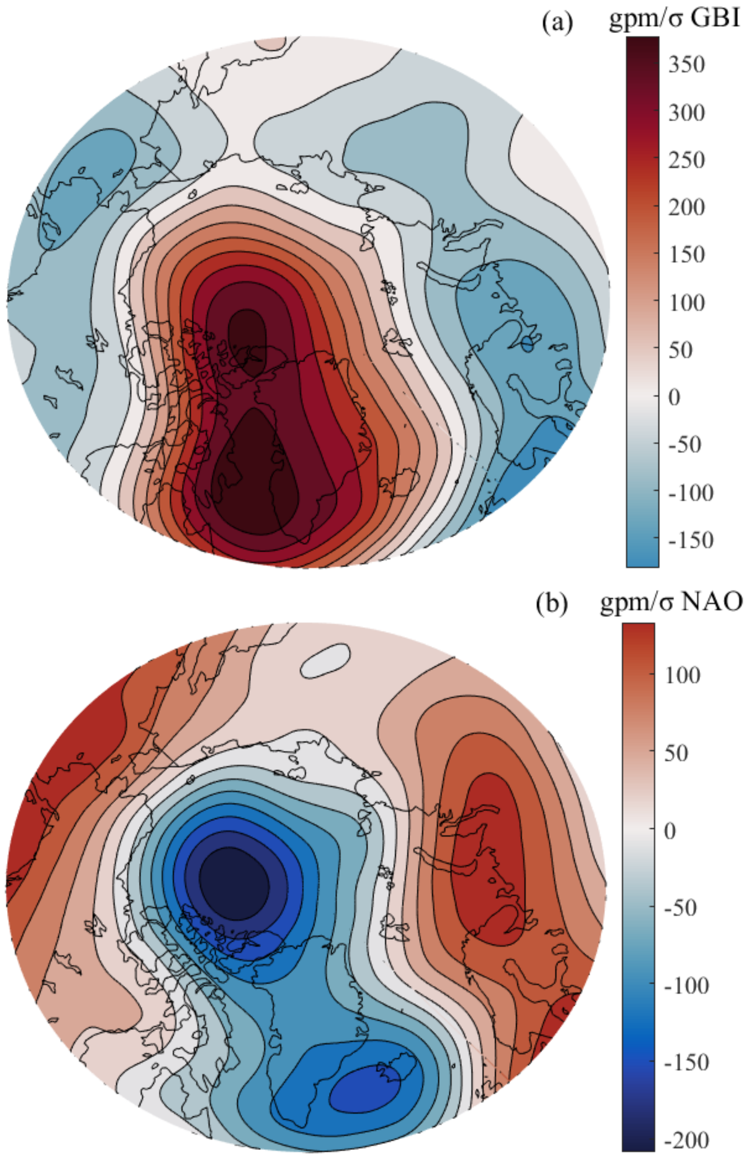
Next we discuss the spatially different response of western GrIS climate and melt to GBI. To that end, Fig. 12 shows maps of the JJA GBI dependency for temperature (Fig.12a) and melt (Fig. 12c) for Greenland and its immediate surroundings using RACMO2. Fig. 13a shows the regional 500 hPa height anomaly from ERA5 associated with variations in GBI.

Figure 12 and 13: missing subplot titles and colorbar labels.

Reply: Combined with the first reviewer's suggestions with the "diverging color ramp", we have changed Figures 12 and 13 as requested.



New Figure 12



New Figure 13

References

Reply: We have added these references.

Charalampidis, C., van As, D., Box, J. E., van den Broeke, M. R., Colgan, L. T., Doyle, S. H., et al. (2015). Changing surface-atmosphere energy exchange and refreezing capacity of the lower accumulation area, West Greenland. The Cryosphere, 9(6), 2163-2181.

Vandecrux, B., Fausto, R. S., Langen, P. L., van As, D., MacFerrin, M., Colgan, W. T., et al. (2018). Drivers of firn density on the Greenland ice sheet revealed by weather station observations and modeling. Journal of Geophysical Research: Earth Surface, 123, 2563-2576.

Favier, V., P. Wagnon, J.P. Chazarin, L. Maisincho, and A. Coudrain (2004), One-year measurements of surface heat budget on the ablation zone of Antizana Glacier 15, Ecuadorian Andes, Journal of Geophysical Research, 109(D18), doi:10.1029/2003 jd004359.

Kuipers Munneke, P., van den Broeke, M. R., Reijmer, C. H., Helsen, M. M., Boot, W., Schneebeli, M., and Steffen, K.: The role of radiation penetration in the energy budget of the snowpack at Summit, Greenland, The Cryosphere, 3, 155-165, https://doi.org/10.5194/tc-3-155-2009, 2009.