We would like first to thanks both reviewers to their useful and constructive comments.

As both reviewers suggest, we propose to simplify our paper

- by removing the distinction between MARv1 and MARv2 (see response #13).
- by putting the sections (3.2 and 4.3) about the seasonality in Supplementary Material.
- by considerably reducing the abstract.
- by focusing more on the aim of this manuscript which is to estimate the GrIS SMB contribution to the sea level rise. For doing this, we have first used MAR projections but the uncertainties are very large following the GCMs used. That is why we have proposed an estimation of the whole ice sheet SMB changes using only GCMs outputs to have an idea of the range of the projections using all the GCMs from CMIP5.

We don't plan to reduce significantly the validation over current climate because this manuscript will serve as reference for the papers using the MAR future projections.

Finally, we are aware that our 3 selected GCMs could be wrong in the other parts of the world and are likely not the best for estimating the amplitude of the global warming following a given scenario. That is why, we have discussed in the manuscript the SMB components vs temperature anomaly to manumit of the problem "when such temperature anomalies could occur ?". In addition, we have made an estimation of the SMB changes projected by all of the available CMIP5 GCMs because it is likely that the CanESM2 and NorESM1-M based projections are out of the range of the CMIP5 multi-model mean. Only the MIROC5-forced ones seem to be more likely and suitable for forcing an ice sheet model because the MIROC5-forced simulation works well over current climate and its future projections is near the CMIP5 ensemble mean.

Our responses (in blue in the text) to the reviewers remarks are numbered to facility the cross-referencing.

Reviewer # 1

The authors address the issue of the future surface mass balance of Greenland through experiments using 2 RCMs forced at the boundary by GCMs from CMIP5. They attempt to determine if the uncertainty in projected SMB can be reduced if only GCMs, which they judge to perform well for the recent past, are used.

The paper represents an incremental advance on Rae et al (2012).
However, the paper is poorly presented both in the objective of the study and the internal logic towards a conclusion. It does not maintain a consistent storyline and so becomes very difficult to read. It would appear that the authors have produced a brain-dump from their RCM simulations and hoped that the mass of information would produce a viable paper.
A major rewrite and rethink of the intended message and its presentation is required.
I have made a number of suggestions as to changing the information content and style, but not wanting to rewrite the paper myself, I hope the authors can apply similar modifications throughout.

Detailed comments

Abstract: The abstract is too complex – it is not necessary to cover every aspect of the paper here. A bad idea to have a list embedded in the abstract – a list which adds nothing to the headline conclusions. Suggest rewriting to express the motivation (missing), method (first sentence), and headline conclusion (lines 21-25) with implications.

1. We will reformulate in depth the abstract according to your suggestions by focusing more on the aim of this paper (cited earlier).

Introduction:
3103-12:14. What about the impact of increased liquid precipitation?

2. It is clear that the increase of rainfall wets the snowpack, decreases the snow albedo and then enhances the melt. But, as discussed in Sec 5.1 in the paper (pg 3121), most of the rainfall increase occurs over bare ice areas in the ablation zone and therefore does not have a consequent impact on the melt because the bare ice albedo is fixed in MAR.

3103-22:23. Redundant. Reduce this sentence to "Mass loss from ice calving is estimated to be roughly that from SMB van den Broeke et al., 2009; Rignot et al., 2011)."

3. thanks, we will reformulate our sentence according to your suggestion.

3103-22:24 Suggest changing this to "The dynamical response of the ice sheet to surface melt reaching the bedrock (Zwally et al., 2002) is still uncertain. However, recent observations suggest in is not a significant component of the total uncertainty in future sea level rise (Sundal et al., 2011 Rignot et al., 2011). The acceleration in flow of tidewater glaciers, due to large melting at the calving front, is expected to decline in future as the glaciers retreat above sea level (ref)."

4. Thanks, we will change our sentence according to your suggestion. For the last reference, we will site :

Goelzer, H, P. Huybrechts, J.J. Fürst, M.L. Andersen, T.L. Edwards, X. Fettweis, F.M. Nick, A.J. Payne and S. Shannon (2012), Sensitivity of Greenland ice sheet projections to model formulations, submitted to the Journal of Glaciology.

3103-27. Reword this confusing sentence, perhaps "A preliminary objective, addressed in this study, is to provide the best possible estimate of future surface mass balance and associated surface freshwater runoff Such runoff contributes both to sea level rise and may eventually affect the North Atlantic thermohaline circulation (Swingedouw et al., 2009)."

5. Thanks, we will reformulate our sentence according to your suggestion.

3104-13:15. Although it is true that GCMs do not include snow-firn-ice physics, this is not to computational load. The snow scheme has to work for seasonal snow cover (important for regional water resources) and all vegetation types, it does not follow that the snow schemes are inferior. In addition the ice sheets are often fudged to avoid ocean salinity drift, so providing better understanding of large-scale changes to the global hydrological cycle.

6. We will reformulate our sentence by this: *Moreover, the GCMs usually lack a realistic representation of the snow/firn/ice physics because the same snow scheme has to work for seasonal snow cover, all vegetation types and ice/snow covered areas of all climates while the snow scheme can be tuned for a specific area in a RCM.*

3104-16:17. Rephrase this. RCMs contain neither feedbacks on ice dynamics nor impacts on THC, so they are not 'ideal' tools. Be more specific about the purpose of RCM's in solving overall mass balance 'issues'. As GCM resolution increases (examples in AR5) RCM's will become redundant. You mention in 3106-13:17 that there is qualitatively no difference between 15 & 50 km resolution, so a more consistent storyline is required. . .RCMs can assist the development of GCMs specifically that of snow-firn-ice schemes (actually in CESM for AR5 but now being incorporated in many others).

7. We want not here enter the debate RCM vs GCM.
- For short time future projections (100 yrs) as considered here, potential changes of THC coming from the GrIS melt can be neglected here. For long time future projections, we agree that GCM are better in this case.
- It is true that the feedbacks from the ice dynamics can be significant (evaluated by Goelzer et al. (2012) to be 7-17% of the SLR coming from GrIS); that is why we conclude this paper by saying that we need to couple MAR with an ice sheet model. The new version of MAR will allow to do this with an ice sheet topography and mask able to vary during the simulation.
- We agree that GCM will soon reach the spacial resolution of the RCMs and that, for making future projections, very high resolutions as the RCMs can reach are not needed knowing the uncertainties in the scenario. However, the main difference between a RCM and a GCM is that the physics of a RCM can be developed for a specific area while a GCM should work for all the climates. To our knowledge, it is difficult to have a model which works well everywhere with the same parameters and physics. As you suggest hereafter, while the GCMs chosen in this study (e.g. CanESM2) work well over Greenland, they could fail to simulate the climate at the global scale and therefore the amplitude of their projected changes could be questionable for a given scenario. Other example, although HadGEM2-ES is too warm in summer over Greenland, it is clearly one of the best ones over Europe and at the global scale and therefore, the amplitude of the projected change are likely more reliable. Therefore, a RCM developed for specific climates will be better than any GCM if it aims to simulate the climate at the global scale.

We suggest to add a small paragraph summarizing these issues in the manuscript.

3104-26. But MAR was used in Rae et al (2012), so this study is not uniquely different.

8. Yes, MAR is indeed used in this paper. But we refer here to the other RCMs (HadRM3P and HIRHAM5) used in Rae et al (2012) and by comparison with MAR and RACMO2, the authors found that the RCM inter-comparison highlights the importance of using a detailed snow physics scheme, especially regarding the representations of albedo and meltwater refreezing.

3105-19:22. Rearrange this sentence to avoid repeated use of the word 'module'.

9. OK we will do

What is meant by 'integrated surface albedo' is it an emergent characteristic of the proceeding components (for example using a delta-eddington scattering model) or is it merely an empirically tuned expression of grain size and wetness?

10. The surface albedo is computed by the CROCUS formulation using grain size and form. This formulation has empirically been tuned with snow measurements from the "Col de Porte" in the French Alps. We refer to Brun et al. (1992)[1] for more details about this formulation.

3105-22:23. Remove reference to blowing snow model. That it exists is not relevant and in any case is negligible (see 3103-9).

11. Indeed. We will remove this sentence.

3105-23:26. Simplify. Suggest: "SISVAT does not include a 3D ice sheet model and consequently Greenland maintains a fixed height and extent through the simulations"

12. Thanks for the suggestion.

3106-7: Either explain here why orographic smoothing is needed, or don't mention it at all (it is irrelevant to the processes described in the paper).

13. An orographic smoothing is needed for improving the stability of the model in the hilly areas (e.g. a long the south-east coast).

But, with the aim of simplifying our paper, we propose to remove the distinction between MARv1 and MARv2 as you suggested and then to remove this section about the orographic smoothing.

Knowing that differences in ice sheet mask and topography impact the SMB results (as shown by Vernon et al., 2012[2]), we suggest to cross-calibrate the MARv1 and MARv2 results by using MARv1_ERA-40 vs MARv2_ERA-40 over 1960-1999. A constant corrective factor will be applied to the MARv1 results (i.e. MAR forced by the ICE2SEA GCM) given in absolute values (i.e. Table 2, Fig 1 and Fig 3 and Fig 5a). A corrective factor will use the differences between MARv1_ERA-40 vs MARv2_ERA-40 in 1D and in 2D.

We just suggest to explain that the ICE2SEA-forced MARv1 outputs has been cross-calibrated with MAR_ERA-40 to be coherent with the other MARv2 results presented here.

3106-15:17. Only one resolution is used throughout the paper, why mention that the parameterisations are resolution independent. There is no reason to suppose that any of the surface components outlined on 3105 should be scale dependent.. All this

1 Brun, E., David, P., Sudul, M., and Brunot, G.: A numerical model to simulate snowcover stratigraphy for operational avalanche forecasting, J. Glaciol., 38, 13–22, 1992.

2 Vernon, C. L., Bamber, J. L., Box, J. E., van den Broeke, M. R., Fettweis, X., Hanna, E., and Huybrechts, P.: Surface mass balance model intercomparison for the Greenland ice sheet, The Cryosphere Discuss., 6, 3999-4036, doi:10.5194/tcd-6-3999-2012, 2012.

statement tells us is that you don't need high resolution to model the ice sheet SMB – perhaps a subject to be presented in the conclusions.

14. Franco et al. (2012) show indeed that the trend of the SMB components simulated by MAR forced by ERA-INTERIM over 1990-2011 are independent of the resolution used in the model. But, in absolute values, the results are different because MAR simulates few more precipitation at lower resolution. It is likely a model artefact. In addition, the resolution impacts also the representation of the narrow ablation zone (length < 100km) and then the amount of melt. In addition, Franco et al. (2012) show that a resolution of minimum 25 km is needed for resolving the main mass balance zones along K-transect and more generally, a resolution of 10-15 km is needed to resolve the steep 5 km slopes in the vicinity of the ice sheet margin. MAR has not been tested at lower resolution than 50 km but we think that in this case we should fail to represent so accurately the ablation zone.

The aim of our sentence is just to explain that our future projections are likely independent of the resolution used when they are integrated over the whole ice sheet. It is not the case locally if we want to force an ice sheet model. But nevertheless, it is clear that, in view of the large uncertainties, performing future projections at very high resolution (1-5km) is not relevant until now, even with the aim of forcing an ice sheet model.

3107-5-15. Rearrange this section to start with verification of MAR-erainterim for the various processes which then supports its use as a reference simulation throughout the analysis.

15. OK good suggestion.

A caveat that the relatively short ERA-interim timeseries could contain significant decadal variability with cannot be temporally matched to GCM forced simulations (which will all have different phases of the decadal variability over this period.). One way to reduce this 'observational bias' is to look at the smoothed NAO index for ERA and GCM to determine if one would expect a difference from a circulation perspective.

16. The comparison of the smoothed NAO index over 1960-2011 is given in Fettweis et al. (2012)[3]. Indeed, after 2000, the ERA-INTERIM shows a decrease of the NAO index while no GCM simulates such a decrease which explains why the reference period is 1980-1999 (when the NAO is quite stable) and not 1980-2010.

In addition, there is below the same figure as Fig 2 but over 40 years !

3 Fettweis, X., Hanna, E., Lang, C., Belleflamme, A., Erpicum, M., and Gallée, H.: Brief communication "Important role of the mid-tropospheric atmospheric circulation in the recent surface melt increase over the Greenland ice sheet", The Cryosphere Discuss., 6, 4101-4122, doi:10.5194/tcd-6-4101-2012, 2012.

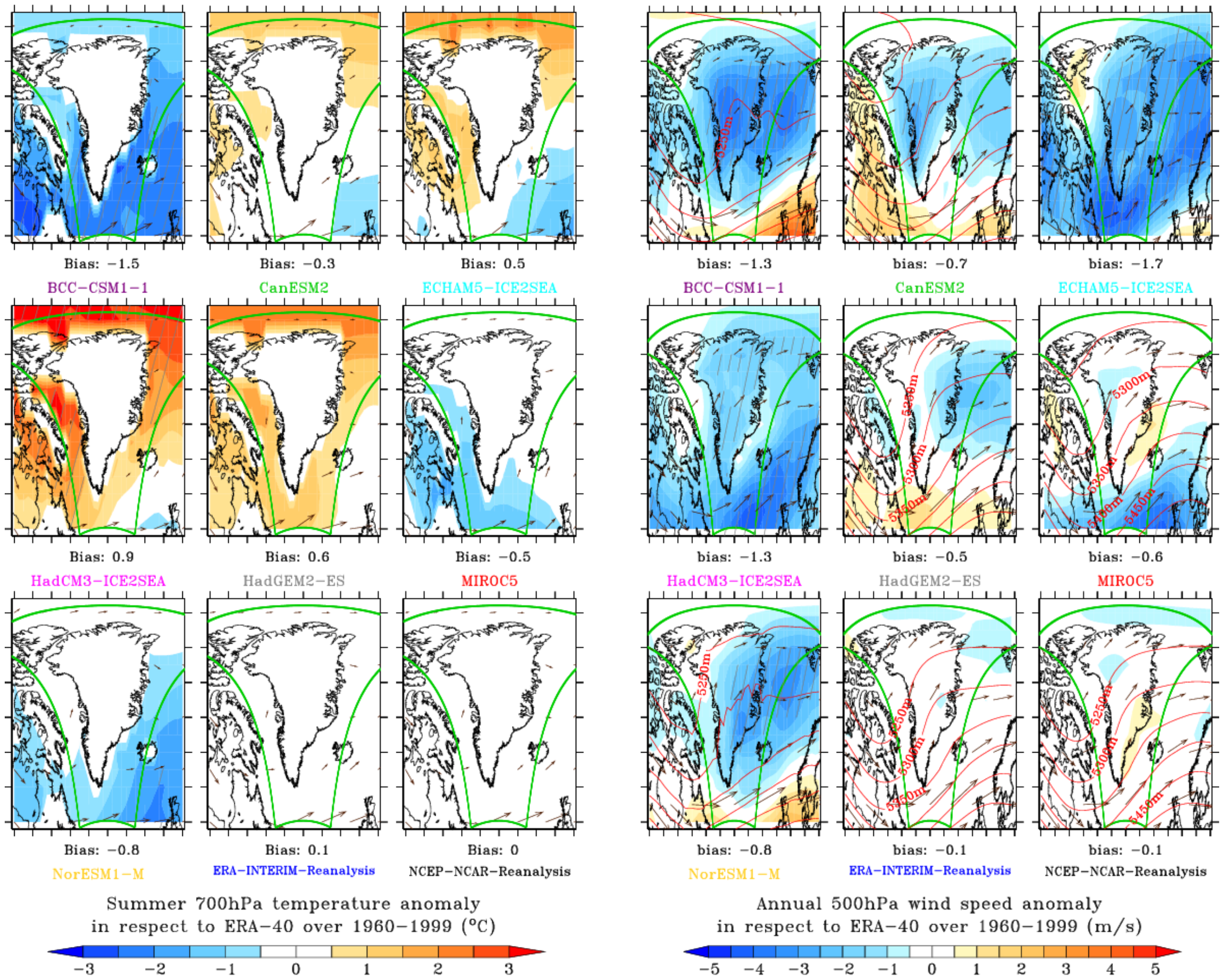

Fig. A Same figure as Fig 2 but for GCM vs ERA-40 over 1960-1999.

This does not change the comparison showing that using 20 yr from ERA-INTERIM is enough here for evaluating the current climate over Greenland.

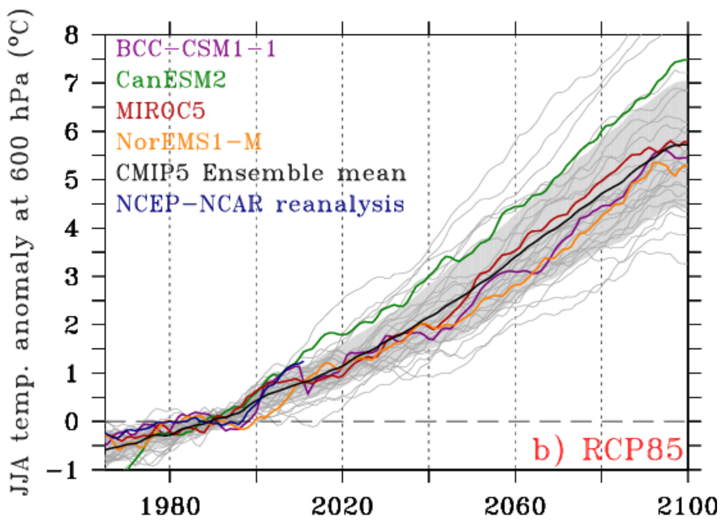

Fig B: Same as Fig 4b but where the period has been extended to 1965-2100.

As you see, from the end of 70's to almost the end of 90's, the summer climate from both reanalyses and GCMs is very stable. Therefore, a comparison over 1980-1999 is enough here.

Analysing a GCM just over the very small area of Greenland for just 20 years is no indication at all that the GCM will perform well in the future. The GCM needs to verify well globally to have any confidence that patterns of change – teleconnections – are valid. The models which 'perform well' over Greenland are actually rather poor globally. For example, CanESM2 depicts that we have already lost the Arctic summer sea ice.

17. Indeed, a GCM which is good over current climate does not mean that it will perform well in future. But, knowing that no GCM project changes in general circulation over Greenland (Belleflamme et al., 2012), we can assume that their general circulation should also be good in future climates over Greenland.

About the amplitude of the warming, it is clear that some of our selected GCMs can be wrong because they fail to represent the climate at the global scale. For example, CanESM2 projects the highest temperature increase likely because of the total disappearance of its summer Arctic sea ice. That is why, we discuss in the manuscript the SMB components vs temperature anomaly to manumit of the problem "when such temperature anomalies could occur ?". In addition, we have made an estimation of the SMB changes projected by all of the available CMIP5 GCMs because we are aware that the CanESM2 and NorESM1-M based projections are out of the range of the CMIP5 multi-model mean. Only the MIROC5-forced ones seem to be more likely.

3107-13:16. The pseudo – endorsement of MAR from Rae et al., against just 2 other RCMs, is unnecessary and should be deleted. No mention at all is required of RACMO2 here. Reference it when relevant later. I suggest removing all text between 'In addition . . .. at a resolution of 11km" to prevent drift in the storyline of the paper.

18. OK we will remove the reference to the RCMs intercomparison of Rae et al. (2012). In addition, saying that RACMO2 has been run at 11 km is indeed not relevant here but using a different ice sheet mask is important following Vernon et al. (2012) because this explains in part, the differences between the MAR and RACMO2 time series.

3107-25:26. Not clear why ECMWF-forced simulations are mentioned in a paragraph about future projections. Rearrange.

19. This sentence reminds that SST and SIC in MAR comes from the forcing. It is just to say that the same setup than the ECMWF-forced MAR simulations is used here. We do not think that this sentence should disturb the reader.

3108-2. A radiative forcing of 4.5 Wm2 at 2100 is of course the definition of the RCP.

20. It is clear indeed but we think that the definitions of the new RCM scenarios are not clear for all

readers and that a reminder is useful here.

3108-5. Replace ‘pessimistic’ with the less judgemental ‘high-end’. These are “representative pathways” not “predicted pathways”.

21. OK for replacing "pessimistic".

3108-9:11. It would have been more interesting to include RCP 2.6 as this is a strong mitigation pathway and might indicate that such a policy trajectory would prevent a –ve SMB.

22. MAR has only been forced by NorESM1-M for RCP2.6. The results of this simulation (presented in Table 3) are very comparable with the NorESM1-M RCP4.5 forced MAR simulation. We could make the same figure as Fig. 8 for RCP 26 and RCP 60 but the problem is that, for a lot of GCMs, only RCP45 and RCP85 scenarios are available. That is why we have chosen to discuss here only RCP45 and RCP85.

3108-17. What about the 60’s & 70’s is comparable to 1980-1999?

23. The SMB and in particular the summer melt was stable from the 60's to the end of the 90's. This is shown in the brief communication of Fettweis et al. (2012). We will add this reference in the text as well as we will refer to SMB and melt by saying that the 60’s & 70’s is comparable … .

3108-20:21. Why is this a problem? Would this not help verify the icesheet behaviour with warming (partially related to a change in NAO).

24. Knowing that no GCM projects changes in the general circulation (Belleflamme et al., 2012[4]; Fettweis et al., 2012), the warmer climate of the 2000's decade resulting from changes in NAO can not be considered as representative of the average present climate over Greenland. As you evoked earlier, it is better to make the comparison on a period when the climate is stable. However, when we discuss the inter-annual variability, the whole period 1980-2011 is considered.

3108-22:26. A good representation of climate over Greenland is no guarantee at all that the large-scale circulation changes with global warming will be any better than a random guess for the future. As mentioned before, regional patterns will change depending on ocean heat uptake, over turning circulation, cloud phase change etc. which means that present day climate of a small region like Greenland will sufferer from large multi-decadal variability. This paragraph needs to qualify statements within this context.

25. Indeed, we will add some sentences explaining that the amplitude of the warming for a given scenario could be wrong even if the model performs well over Greenland for current climate because the model sensitivity depends more of the simulation of the climate at the global scale and

4 Belleflamme A., Fettweis X., Lang C. and Erpicum M: Current and future atmospheric circulation at 500 hPa over Greenland simulated by the CMIP3 and CMIP5 global models. Climate Dynamics, doi: 10.1007/s00382-012-1538-2, 2012.

not at the local scale. But, as said earlier, the aim of this paper is not to select the GCMs projecting the most likely future.

3109-2. Replace ‘well-known’ with a reference.

26. OK

3109-5. ‘..are obviously..’ only if they lie outside the multi-decadal variability – either inferred from the GCM control simulation or an analysis of multiple reanalysis from 1950’s.

27. Yes, indeed but the period 1980-1999 was chosen because the climate was stable in both GCM and reanalysis. It is clear that a comparison over 1970-1999 would be better but ERA-INTERIM is only available after 1979 and the hadCM3 and ECHM5 outputs used here are available since 1980. However, Fig A (See response #16) comparing GCMs over 1960-1999 with ERA-40 shows the same biases than the GCMs vs ERA-INTERIM over 1980-1999. In addition, Belleflamme et al. (2012) comparing the general circulation over 1961-1990 found the same biases in the general circulation.

3109-15. Remove ‘and 600 hPa’ as this is not relevant here.

28. OK

3109-20:29. The separation of windspeed and direction is strange as advection of warm air depends on the source of the air and so wind direction is important as well. The circulation over Greenland is dependent in the position and strength of the polar vortex in the model. It also depends on the Froude number of its interaction over the orography which is RCM dependent.

29. OK but the wind direction is well discussed in the next paragraph via the Z500 pattern. In addition, for a more detailed evaluation of the GCM general circulation, we will refer to Belleflamme et al. (2012) which is now published in Climate Dynamics.

3110-8. Remove ‘indeed’, its use is misplaced here.

30. OK, thanks

3110-11:29. The formulation of the argument is muddled. I suggest restructuring such that you are assessing a number of CMIP5 models. It should not be relevant that they are meteorologically the ‘best’. As detailed below, this is no guarantee that they will perform well in the future scenarios.
I am unable to assess the methodology used to evaluate the circulation regime (paper submitted). I presume it uses some form of self-organised mapping technique (cluster approach). To restrict this to Greenland itself assumes that Greenland exists without teleconnections to the rest of the global system. That some models cluster well against a short term reanalysis, only indicates that the driving GCMs happen to have an NAO and PDO aligned with the present day. Since the climate models are free running, this is unlikely to be the case for most of them. This does not indicate that they would

perform better in the future simulations. The models selected here are particularly bad on the global scale which suggests that there may be serious issues with their parameterisations or process physics.

31. This issue has already been discussed in our previous responses (#17). We suggest indeed to remind in the text that a GCM which is good over Greenland for current climates does not mean that its sensibility to a GHG increase is good and that it is good for the other parts of the world. It is very likely that the GCM chosen here are not the best ones at the global scale. For example, HadGEM2-ES is clearly better than CanESM2, NorESM1-M and MIROC5 over Europe.

3111-3:8. Restructure such that the purpose of the table/figure is defined before it is referenced in the text. E.g. 'The SBM describes the integral response of the ice sheet to the climatic forcing. To identify differences in the forcing we break the SMB down into its components (Table 2).' Table 2. Remove reference to blowing snow in caption.

32. OK, we will reformulate our sentence.

3111-3:8. Describe how the errors on Table 2 are derived and their significance (2 sigma?). I assume these are not just interannual variability, but rather standard errors of the means.

33. The errors are well the inter-annual variability (i.e. the standard deviation). We will update the legend of the table for explaining this.

3111-10:16. What is the purpose of showing MARv1, if the difference between v1 and v2 is just the ice mask? The less irrelevant data you present, the simpler your argument, and the easier the paper is to read.

34. See our response #13 earlier about this.

3111-17:21.Why? Is this due to a misplaced gulf-stream or smoothed orography, ill positioned centre of the NAO, spectral model? I know you do not want to get into adetailed analyses of each GCM, but what is the significance of this description?

35. This sentence just explains why MAR forced by BCC-CSM1 underestimates the snowfall in respect to MAR forced by ERA-INTERIM. It is the same explanation for MAR forced by HadCM3 and ECHAM5 which also underestimates the precipitation along the south-eastern coast. Why these GCMs fail to represent the general circulation over Greenland is not the aim of this paper. These GCMs have also large temperature anomalies in respect to the reanalyses. Therefore, if the temperatures are wrong, the general circulation is wrong, … In brief, a detailed analysis of theses GCMs at the global scale is needed here but this is outside of the focus of this paper.

3111-27. Suggest changing to 'In addition to its impact on SMB, a low winter snowfall results in an earlier exposure of bare ice during summer melt, and since ice has a lower albedo results in a higher ablation (Mote, 2003; Tedesco et al., 2011).

36. OK thanks.

3112-4:'south-westerly flow'

37. OK thanks.

3111-3113. Much of this descriptive reasoning for GCM ‘biases’ is by inference. Showing GCM moisture and heat convergence may be a more convincing approach.

38. Indeed, but the heat/moisture convergence is not available in the CMIP5 data base and we prefer to use outputs coming directly from GCMs for making the inter-comparison. We have evaluated the specific humidity at 600 hPa and 850 hPa and all the GCM seems to be good.

Also the annual mean circulation is depicted in fig 2, but you refer to winter precipitation, the circulation patterns for which cannot be inferred from an annual mean.

39. The pattern of the annual mean general circulation is quasi identical to the one from the winter/summer mean general circulation. That is why an evaluation at the annual scale is enough here. We will add a sentence to specify this.

Reference to ‘common’ circulation features in each section of the description is repetitive. Sat the start of this section, a description of the climatology (from ERA-Interim) is needed (perhaps move the section on this from earlier in the paper). This description would be sufficient to set the scene when describing deviations in the GCMs from the reanalysis.

40. Good suggestion. We will do this.

3113-23. Why is the mean annual cycle important? What does it tell us amount the system? You are just throwing a lot of data at the reader and providing no context.

41. Indeed, the mean annual cycle is less important in respect to the aims of this paper. That is why, we suggest to put Section 3.2 and 4.3 in Supplement.

3114-15. Poor structure + needs context. I suggest .. ‘An adequate depiction of the model observed interannual variability is an indicator that it will perform well under future simulations. The variability arises due to changes in the storm tracks, the NAO, the westerly winds, and sea ice conditions, which may combine to generate decadal variavility. However, alignment of these components in GCMs is not likely to be the same as those in the real world. Consequently, we cannot expect temporal coincidence of any observed changes to those in the GCMs. The summer air temperature anomaly at 600 hPa may be considered a proxy for the evolution of surface melt (Fettweis et al., 2012).

42. Thanks for this suggestion.

3115-18 ‘..temperature sensitivity to GHG increase. . .’
3115-19 ‘’. . .are amongst the GCMs with the highest climate sensitivities’

43. Thanks.

3115-16:26. As noted above, the observed acceleration of melt is not neccesarilly a direct response to global warming. There are a number of 'natural' multi-decadal oscillations that are combining in addition to global warming. It is thus not unexpected that the models will have a mixed response over a 30 year period.

44. OK

3117-6:11. This section should be moved to the introduction of Section 3. It is a characteristic of climate models – that they will not produce a temporally equivalent change in a natural (or even forced) oscillation. However, they will still simulate natural oscillations. Thus the conclusion that one should use the 1980-1999 section against GCMs is invalid, as the GCM is just as likely to have entered the 2000-2011 part of its equivalent oscillation earlier as later. The GCM can still be good yet disagree with 1980-1999 due to natural variability. The conclusion should be that 20 years is not a sufficient time to access the climatology of a GCM!

45. We agree with this issue if the period chosen for comparing GCMs and reanalyses is a transitory period between two climates. Nevertheless, according to Fig. A (response #16) and Fig.2 (in the manuscript), comparing GCMs over 40yrs (1960-1999) or 20 yr (1980-1999) does not change the comparison (e.g. BCC-CSM1 is still to cold in summer, the general circulation from HadGEM2-ES is still the best, ...) because the climate from both GCMs and reanalyses are quite stable over 1960-1999 in respect to the 21th century.

Bellow, we have compared HadGEM2-ES from different periods in respect to ERA-40 over 1980-1999. It is still to warm in summer while the biases are reduced if we compared 1960-1969 from HadGEM2-ES with 1980-1999 from reanalysis.

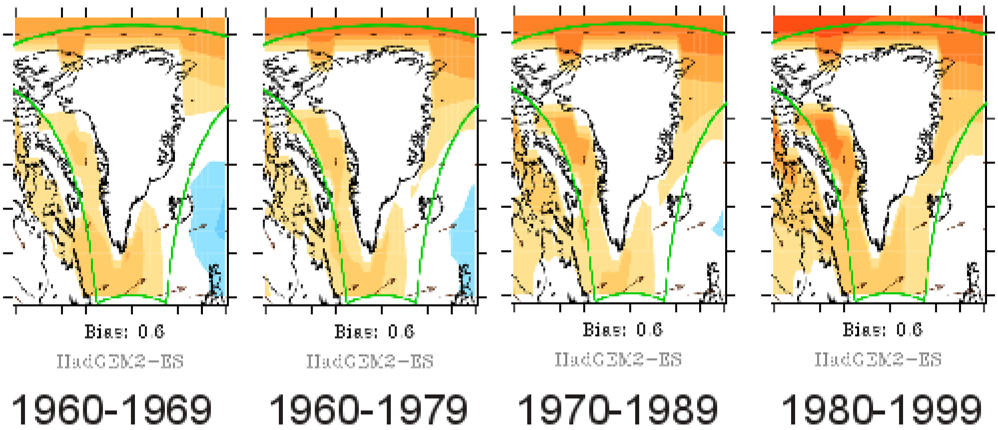

Fig. C Same as Fig2a but for HadGEM2-ES taken over different periods in respect to ERA-40 over 1980-1999.

If we compare the time series of JJA T600 over 1965-2010 in absolute values (see Fig. D bellow), we see well that hadGEM2-ES is too warm during the whole shown period and that BCC-CSM1 is too cold.

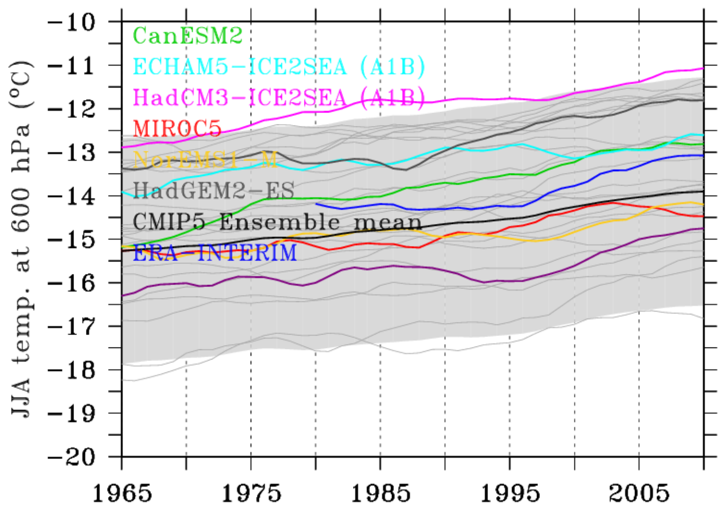

Fig D. Time series over 1965-2010 of the JJA T600 in absolute value (not in anomaly here).

Therefore making comparison over 1980-1999 is enough here in a first approximation and changing the reference period does not change the comparison. We use 1980-1999 for having the same reference period for all the anomalies shown here. It is clear that comparing over 1980-2010 is more litigious because the reanalyses show a strong temperature increase not linked to the global warming but to the NAO.

We suggest to add in Supplementary Material Fig. D and the equivalent of Fig 2. and Fig. 3 but with a comparison over 1960-1999 and over 1960-1990 using ERA-40 as reference instead of ERA-INTERIM and to explain in the manuscript a potential issue coming from a comparison over transitory periods as the 2000's. We will also add a reference to Belleflamme et al. (2012) who compare T700 from GCMs over 1960-1990.

3118-10:15 I don’t understand this paragraph – rephrase. Surely any bias in 1990 mass balance is small compared with the differences by 2100 (as per your comments on 3117-20:26)?

46. If we want to estimate when the SMB integrated over the whole ice sheet becomes negative, we can not use the SMB projections in absolute value (Fig5a) but well in relative value (i.e. anomalies in respect to 1980-1999, see Fig. 5b). Because, for MAR forced by HadCM3, a negative SMB in absolute value should already occur at the end of the 90's! We will rephrase this paragraph.

3118-14. Replace “most pessimistic simulations” with “high CO_2 emission scenarios”.

47. OK thanks

3118-17:18. The argument is messy since you are not doing attribution studies (e.g. keeping sea ice fixed). Thus it is better to state what you would expect to see then determine what the models indicate. For example, I would expect to see a shift in the zonal winds (which can be tested in the forcing fields) and consequently the pattern of precip. Try to tie the changes to the boundary conditions as you did on 3109. This will tie the elements of the paper together and make a tighter storyline.

48. We assume that the reviewer refer to 3118-24-25 about the impact of the sea ice decrease to the precipitation increase. We have tested MAR over current climate (forced by ERA-INTERIM) by decreasing the sea ice cover and we observe well a small increase of precipitation due to an enhanced moisture evaporation above open water pixels.

3119-18:22. These suggestions are testable. Do the GCMs with the highest sea ice reduction result in the largest winter warming?

49. Yes, the highest warming occurs well for CanESM2 which projects the highest sea ice decrease in our simulations.

Do the simulations with the greatest change in the ELA result in the largest August T anomaly?

50. Yes, it is well the case because a larger August T anomaly (in respect to other months) occurs well if we compare RCP85 and RCP45 using the same GCM as forcing.

It should be noted that we suggest to put this section in Supplementary Material for simplifying our paper.

3119-23:28. Rephrase. "However, in these scenarios we see little seasonal change in the components of the SMB. . ..".

51. Thanks.

3120-1. '. . .most of the simulations. . .' Why don't ALL the simulations have snowfall converted to rainfall?

52. For the simulations projecting a low temperature increase (as forced by NorESM1), there is anyway a snowfall increase in summer at the scale of the whole ice sheet because the climate is warmer although a part of the precipitation increase is converted to rainfall. The figure below (Fig. E) shows that a JJA TAS increase > 3°C is needed for having a JJA snowfall decrease.

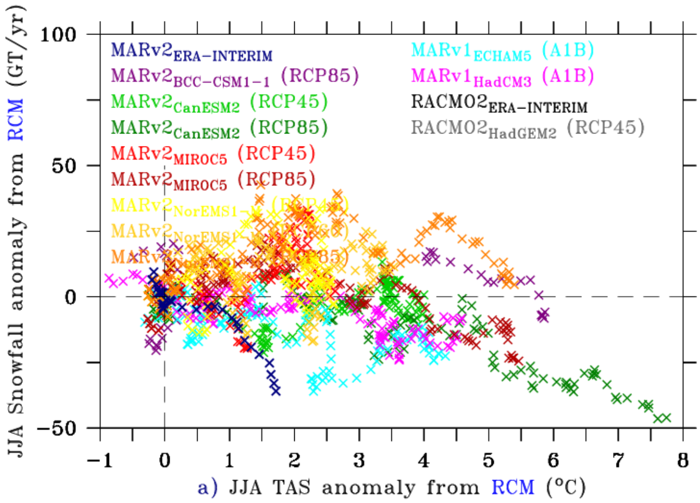

Fig. E. JJA snowfall anomaly vs JJA TAS anomaly.

3120-4:6. Poor sentence structure – suggest "An increase in summer snowfall in the percolation zone would result in a negative albedo feedback (Box et al., 2012. . ..), but no such snowfall change occurs in the MAR simulations.

53. Ok thanks.

3120. What about the variability in summer temperature? If in some models the temperature extremes become greater, with regional warming, then the daily melt events are stronger even if the mean temperature rise is the same. Are there non-linear characteristics such as this which could cause the models to produce different SMB by 2100?

54. It is a very good question. It is clear that an increase of the variability in summer should impact the amount of melt. We have plotted below the JJA temperature variability (i.e. the standard deviation of the daily temperatures) averaged over 1980-1999 and over 2080-2099 using MAR forced by MIROC5 (historical and RCP85). As you see, there is rather a small decrease of the daily variability.

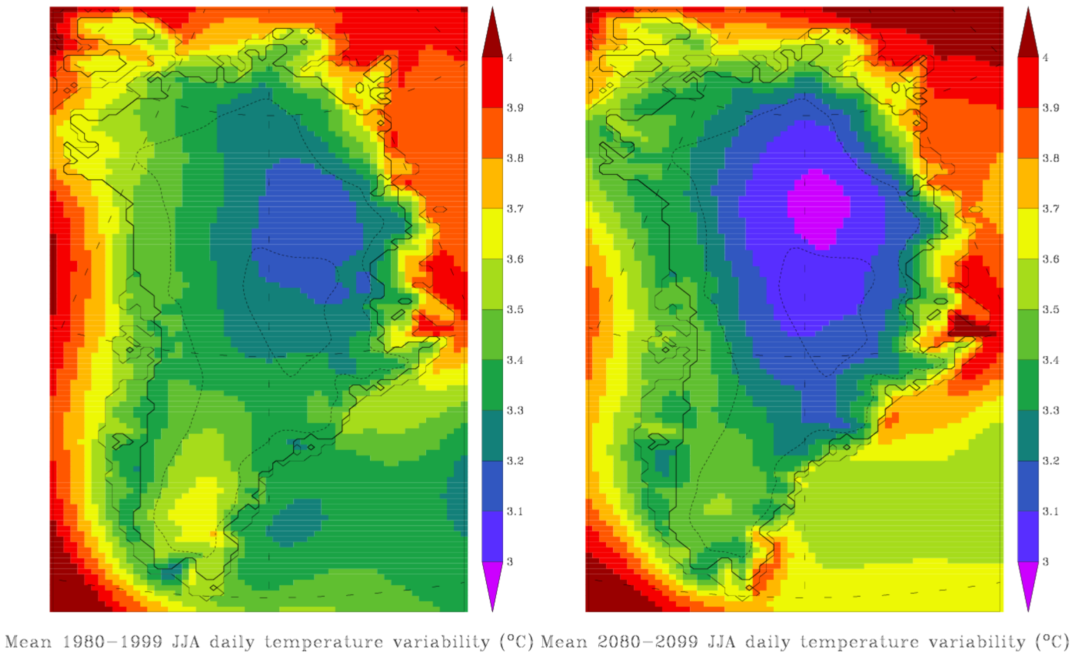

Fig. F JJA temperature variability (i.e. the standard deviation of the daily temperatures) averaged over 1980-1999 and over 2080-2099 simulated by MAR forced by MIROC5.

3121-11:16. Suggest rearranging this to ‘However, rainfall does not contribute significantly to the SMB, occurring mostly over bare ice or saturated surfaces. Since the rain is not considered to transfer sensible heat, it simply runs off in both MAR and RACMO2. Consequently, we only consider meltwater (run-off minus rainfall) hereafter.’ Check that you are consistent in the use of the term ‘meltwater’

55. OK thanks.

3121-17. ..warming climate. . .

56. OK

3121-17:27. This section need to be rearranged and simplified (avoiding repetition). That on general increase in precip with temperature (sensitivity) should be at the start of 5.1. It should then lead into paragraphs on rain and snowfall changes. Note that if RACMO2 has a larger domain than MAR then the RCM physics will in part alter the pick-up of water vapour from the ocean within the domain.

57. Ok we will rearrange this section.
About RACMO2, its integration domain is larger than MAR and therefore, the physics from RACMO vs GCM which forces MAR could impact the amount of evaporated water from ocean in the RACMO domain not included in the MAR domain. However, we have shown that a larger domain with MAR do not impact the amount of precipitation over the ice sheet (Franco et al., 2012)[5].

5 Franco, B., Fettweis, X., Lang, C., and Erpicum, M.: Impact of spatial resolution on the modelling of the Greenland ice sheet surface mass balance between 1990–2010, using the regional climate model MAR, The Cryosphere, 6, 695-711, doi:10.5194/tc-6-695-2012, 2012.

About rainfall, RACMO2 simulates a higher increase of rainfall vs temperature increase in respect to MAR. It is likely that this discrepancy results from differences in the cloud micro-physics used in both RCMs (e.g. the threshold for determining when snowfall becomes rainfall).

3122-1:14. The rational for doing this is poorly explained. GCM circulation changes in the Arctic and North Atlantic during global warming will alter precipitation (water vapour convergence) over Greenland. You seem to be calling this 'inter-annual variability'. Why do you need to do this? What does it add to the paper?

58. The inter-annual variability is well here the variability driven by the general circulation from GCMs.

We need this approximation for estimating SMB changes at the scale of the whole ice sheet using GCMs only. This is currently badly explained at the beginning of Section 5 (3120-25). We will reformulate our introduction of Section 5 and more motivate the aims of this paper.

3122-16:3123-7.. This is a confused section and needs rearrangement/compression. If you have the evidence here then you do not need to refer to the work of others otherwise there is no need to mention it in this paper – it is not new. You have already discussed the bare-ice albedo effect so no need to go over it again. These 27 lines can be compressed to . . . 'The GrIS meltwater flux can be approximated by the JJA TAS (Figure 7c). However, the meltwater run-off rate increases with JJA TAS, which may be interpreted as the extension of the summer ablation zone (Franco et al., 2012b). Refreezing within the snowpack is anticipated to decrease due to the formation of ice lenses which reduce the downward percolation of meltwater '

59. Ok thanks for the suggestion. We will simplify this section.

3123-8:18. This needs simplification. Suggest rather than naming simulations, make a generic statement and refer the reader to the table.

60. Good suggestion. We will do this.

3128-23:29. These highlights are structured poorly. E.g. #1. The key finding is that seasonality does not change, so 1. No increase in the duration of the summer melt season with warming, due to enhanced winter snowfall.

61. We will rephrase #1 according to your suggestion.

3129-9:11. This is quite surprising given the GCMs depict changes to the storm tracks, sea ice and AMOC. It is possible that the standing Rosby wave maintains a fixed local climate over Greenland. (given the small domain used this will result in little change at the boundaries.

62. It is clear that it is quite surprising but Belleflamme et al. (2012) evaluated the GCM-based general circulation at a daily time scale using a circulation type classification and found no change between the occurrence frequencies of the different circulation types simulated by the GCMs over current climate and at the end of this century. There is just a generalised increase of the geopotential height due to a uniform warming of the atmosphere. We refer to Belleflamme et al. (2012) for more

details about this. (see http://www.springerlink.com/content/m765331048634740/)

3129-20:27. I disagree that RCMs are useful for such studies unless they feed back on the general circulation models (where elevation changes have a significant impact on storm tracks, Forne effect, Froude number etc and meltwater affects overturning circulation). The effort should be in full ocean-ice-atmosphere coupling in high res GCMs. This is supported by your earlier statements that RCM resolution had little impact. The study here is supportive of such in that it identifies processes that the SMB of GCMs needs to include.

63. See our response #7 about GCM vs RCM. However, we will "temperate" our conclusion.

Reviewer # 2

General Comments
In this paper, the authors use the MAR model to downscale output from a number of GCMs and compare the results with simulations forced by reanalysis data to assess the capability of the model in making projections of sea level rise in the future. The combination of a number of different forcing GCMs from the CMIP5 database make this a novel piece of work and the thorough analysis of results by the authors helps to unpick the strengths and weaknesses of the different forcing GCMs. The structure of the paper is relatively easy to follow but unfortunately the mass of detail presented obscures the importance of some of the issues raised by this study which would reward further discussion. In addition I miss a strong narrative within the paper in terms of setting out the aims of the study from the outset and then presenting methods and results within that context. The conclusions of this study are scarcely surprising, but the exercise of carrying out these simulations and analysing the strengths and weaknesses of each is valuable in terms of tracing the climate envelope for Greenland and surface mass balance of the ice sheet. A re-write and re-structuring of the paper to reflect these points would be very worthwhile with the emphasis on clarifying and refining the analysis of the results and drawing out the key elements while discarding extraneous and distracting detail that is unnecessary to the narrative.
Specific Comments
I agree with many of the comments expressed in the other review and rather than simply repeating these I add some of my own.

1) The abstract is overly long and should be shortened to include only the main conclusions and a brief overview of the methods.

64. See our response #1

2) On page 3106 lines 13 – 16 indicate that when integrated over the ice sheet the interannual variability of the SMB is not affected by different horizontal resolution. In

this paper only 25km resolution is used for MAR, but it is compared with RACMO at 11 km resolution. A paper by Lucas-Picher et al., 2012 shows that horizontal resolution is important when looking at local and even regional climate in Greenland. This does have important implications on a larger scale since ice sheet models are very sensitive to the distribution of SMB and this would have a consequential effect on ice dynamics which may also feedback on to the SMB. Some discussion of this point would be useful since it is returned to on page 3112, where it is clear that some of the models (e.g. CanESM2) appear to get good answers (when integrated over the whole ice sheet) for the wrong reasons (geographical variability of SMB in the model does not reflect the actual pattern of geographic SMB variability).

65. Good remark. Indeed, at the local scale, it is clear that higher resolution is needed as also shown by Franco et al. (2012)[6]. They showed that a resolution of minimum 25 km is needed for resolving the main mass balance zones along K-transect and more generally, a resolution of 10-15 km is needed to resolve the steep 5 km slopes in the vicinity of the ice sheet margin. This is particularly important for forcing ice sheet models. But, Franco et al. (2012) show that the 15km MAR results can be approximated by the 25 km MAR results and therefore making future projections at 25 km is a good compromise between computer time and estimation of the SMB spatial variability. Among our future projections, only the MIROC5 forced one could be used to force an ice sheet model because there is indeed an error compensation of the SMB at the scale of the whole ice sheet in the CanESM2 forced one. In view of the large uncertainties induced by the GCM-based scenarios, performing future projections at very high resolution (5km) is not relevant until now.
We will more discuss these issues in the manuscript.

3) It is not clear why the particular GCMs were chosen in this study, reading page 3107 and 3110 I am left with the impression that it was largely because the output was in the correct format. Is this correct? How is their ability to simulate general circulation over Greenland assessed? And how does this relate to their abilities to simulate circulation in the rest of the world? It is of course entirely possible that Greenland is well represented in these models for entirely the wrong reasons, so much more information needs to be provided here.

66. We refer to Belleflamme et al. (2012)[7] for a detailed evaluation of the GCM-based general circulation at a daily time scale using a circulation type classification. It is clear that the GCMs chosen are the best ones among the GCMs for which 6 hourly outputs were available at the model vetical levels at the time of writing this paper. We have chosen the model following their ability to simulate the summer temperature (T600) which impacts a lot the melt simulated by MAR and following the conclusions of Belleflamme et al. (2012). In appendix, a comparison of each GCMs is given. In addition to CanESM2, NorESM1 and MIROC5 chosen in this study, BNU-ESM and MPI-ESM-MR should also be good candidate for forcing a RCM over Greenland but 6 hourly outputs of those ones were not available at the time of writing this paper.

About the GCM at the global scale, we refer to our responses #17 and #31. It is clear that some GCMs can be good over Greenland for entirely wrong reasons.

---

6 Franco, B., Fettweis, X., Lang, C., and Erpicum, M.: Impact of spatial resolution on the modelling of the Greenland ice sheet surface mass balance between 1990–2010, using the regional climate model MAR, The Cryosphere, 6, 695-711, doi:10.5194/tc-6-695-2012, 2012.

7 http://www.springerlink.com/content/m765331048634740/

4) Also on page 3107, the use of the MAR ERA-Interim simulation as a reference dataset is justified as ‘having been successfully evaluated against in-situ weather station data’. This is a bit vague and some detail on what MAR does and does not represent well would be welcome, for instance are there biases in seasonal temperatures? How does this affect how we interpret the simulations forced by GCMs compared to the reanalysis?

67. We think that the manuscript is already enough long for including a short evaluation of MAR forced by ERA-INTERM. We think that references to previous "validation" papers are enough here.

While MAR is known to be one of the best RCMs over Greenland, it is far away to be perfect as each model. It is for example too cold in winter and overestimates precipitation in the interior of the ice sheet. Knowing that these biases are identical if MAR is forced by ERA-40 or ERA-INTERIM, we can reasonably assume that the GCM-forced simulations are also affected by these same MAR biases and therefore, that these biases do not affect the comparison here over current climate because MAR is compared to MAR. For future projections, it is likely that these biases could have different sensitivity in a warmer climates but it is difficult to evaluate. However, these biases are acceptable over current climate (i.e. below 1 standard deviation), therefore we can assume that they do not affect a lot the future projections.

5) On page 3108 the period 2000 – 2010 is explicitly excluded because it is a warmer decade than preceding ones. It would be interesting to see output from this period though as it is likely to be more representative of the future climate in Greenland and could have some interesting implications for the evolution of the ice sheet.

68. We think that the comparison must be done when the climate is stable as suggested by reviewer #1. As you see in our response #16 and #45, extending to the past the period of the comparison does not change the results.

The 2000-2010 period is a transitory period with a changing climate associated to changes in the NAO. As discussed in Fettweis et al. (2012)[8], the recent warm summers over Greenland can not be considered as a long term climate warming (as simulated by the GCMs) but are more rather a consequence of the NAO variability impacting the atmospheric heat transport. Therefore, knowing that the GCM simulates an average climate and do not project consequent changes in NAO through this century, comparing the GCM over the last decade when the climate is changing due to the more anticyclonic conditions than normal (i.e. a warmer and drier climate) is less relevant we think.

6) The decreasing SMB observed over the last 10 years is problematic for this study in the way it is presented here since it appears to show that the models are only capable of simulating one phase of the natural variability of the system but not others. This is however, largely due to the short period of the simulation. In the light of recent (and still tentative) work suggesting that loss of Arctic sea ice cover may lead to changes in the dominant atmospheric circulation, this point needs some further discussion.

8 Fettweis, X., Hanna, E., Lang, C., Belleflamme, A., Erpicum, M., and Gallée, H.: Brief communication "Important role of the mid-tropospheric atmospheric circulation in the recent surface melt increase over the Greenland ice sheet", The Cryosphere Discuss., 6, 4101-4122, doi:10.5194/tcd-6-4101-2012, 2012.

69. Since the GCMs does not project consequent changes in the general circulation (Belleflamme et al., 2012) and that the recent SMB decrease is due to changes in general circulation (Fettweis et al., 2012), it is normal that the GCMs fail to simulate the current observed trend. We will more insist on this problem in the manuscript

In fact a better way of presenting this may be to simply state that 20 years is too short to assess the climatology of a GCM.

70. As you see in our response #16 and #45, extending to the past or future the period of the GCM evaluation does not change the results.

In this context, the performance of a GCM against the 1980-1999 period as opposed to the later 2000 – 2010 period is more or less meaningless since both should be well within the bounds of natural variability but the model is not in the correct mode at the correct tome.

71. Indeed, according to response #69, comparing 1980-1999 vs 2000-2012 is less relevant and we will be more prudent in the manuscript by discussing this issue.

Even more difficult, just because it performs well against present day data does not mean it will continue to do so. This point should be made much more prominently, especially given the conclusion in section 3.4 that none of the GCMs can adequately reproduce all aspects of current climate in Greenland.

72. See our response #17.

7) Section 5.2 is very confusing and difficult to follow; I suggest simplifying and rewriting this entirely.

73. OK we will simplify this one according to the remarks from reviewer #1.

8) The figures are very small and the text practically unreadable in some of them, if possible it would be nice to have them enlarged and put on separate pages.

74. A part of this problem is due to the "landscape" presentation of TCD. The figures have been made for a "portrait" presentation as TC. We will enlarge them if they are too small.

Technical Corrections

Given that most of the authors are writing in a non-native language, there are few major errors in this paper but I would like (as a native speaker) to give a few comments on style.

1) The text is made unnecessarily complex and difficult to read by the use of multiple nouns or adjectives before a noun in a string. While this is not necessarily incorrect use of English, it is not easy to read and makes it harder to grasp the sense of some

of the sentences. For example line 1 on page 3105 reads “which aims to improve the projections of the land ice melt contribution to future sea level rise”. It would be easier to follow this clause if it were written “ which aims to improve projections of future sea level rise due to the contribution of melt from land based ice”

75. Thanks for the suggestion.

2) A number of prepositions are incorrectly applied throughout the text (I realise that this is a very difficult one to get right when writing in a foreign language!), for example, line 19 on p. 3105 should read “it consists of”, I suggest having the manuscript proof read through by a native speaker to catch these minor errors.

76. The revision will be reread by a native speaker. In addition, it should be noted that TC offers now copy editing of all TCD papers.

3) It’s a personal preference but I dislike seeing “impact” used as a verb, (for example, line 19 on page 3103 “Beside impacting surface processes...”). Verbs such as ‘to affect’, ‘to alter’ or’ to vary’ can be used instead.

77. Thanks for the suggestions.

4) Typos on page 3108 the 4.5 wm-2 and 8.5 wm-2 units should be W m -2

78. Thanks.