

Supplement to author comment on RC2 'Review 2' by Anonymous Referee #2

The manuscript has been revised, incorporating changes from all Interactive Discussion comments. Versions of the manuscript and supplement with tracked changes are included alongside this document. Major changes to the manuscript include:

1. What was Figure 3 has been moved to supplementary material. Two extra stations have been added, and lines made clearer.
2. Two new datasets have been added to the analysis: Box (2013) data, and MAR v3.5.2 forced by ERA-20C.
3. Time series of summer ice-sheet average surface air temperature have been added as a new panel in what was Figure 7 (now Figure 6).
4. Time series of annual surface air temperature at several long running coastal stations have been added to supplementary material.
5. Maps of grid point trends have been added to supplementary material.
6. Annual mean of monthly bias and monthly mean absolute error has been added in a new table.

Point-by-point responses to referee comments are given below.

Referee's comments are reproduced in black.

Authors' responses are in blue.

Interactive comment on "Evaluation of Greenland near surface air temperature datasets" by J. E. Jack Reeves Eyre and Xubin Zeng

Anonymous Referee #2

Received and published: 4 February 2017

This is a timely and fairly novel comparison of Greenland near-surface air temperature (SAT) bias trends between various reanalysis/regional climate model and instrumental based gridded SAT products and in situ meteorological SAT data, and an analysis of Greenland SAT trends over time since 1900.

The paper is clear and well-written, and makes some interesting and useful conclusions about the likely accuracy and reliability of the various gridded SAT products that are analysed. It provides a valuable contribution to the relatively limited literature base on this topic, especially given the recent plethora of available reanalysis and regional climate model products for Greenland (including several products spanning the first half of the twentieth century).

I would like to have seen a bit more spatial analysis of Greenland SAT trends, e.g. gridded maps of trends from several different datasets, showing how the trends spatially vary, where they are most significant and reliable for different time periods, and summarising the differences and explanations in these.

We thank the referee for this recommendation and include maps of trends in the supplementary material.

page 2, line 3: I don't follow why "the benefit for SAT [of high resolution] is less clear [than for SMB]" Why should this be the case, given that melt and runoff strongly correlate with SAT which is much more directly elevation-dependent than accumulation?

Our point here is that, because SAT is strongly elevation-dependent (as are melt and runoff, as you point out), the benefit of a high resolution model over a low resolution model with post-processing (e.g., the elevation corrections used in Hanna et al. (2011), Lucas-Picher et al. (2012) and our study) is not clear. On the other hand, for precipitation (and therefore SMB), high resolution modelling appears to lead to improvement over low resolution modeling. This has been clarified in the text.

p.2, l.17: "The LATTER SAT reconstruction was compared..."

This sentence has been clarified.

p.2, l.29 remove comma after "SAT observations".

Done.

p.5, l.9 How were the 31 model configurations selected from the CMIP5 ensemble?

The model configurations were chosen based on availability of necessary data: SAT from historical runs and ice fraction fields. This has been clarified in the manuscript.

p.5, l.20: add that 20CR also has a positive bias in the central-most regions.

Done.

p.6, l.5 "MERRA2 closely matches the observed seasonal cycle at all stations" - does MERRA2 assimilate data from these meteorological stations and is this therefore a non-independent comparison and unsurprising result?

To the best of our knowledge, the MERRA2 reanalysis does not assimilate SAT from any land stations (in fact, of the reanalyses used here, only ERA-Interim assimilates SAT, which it uses to update surface properties as part of its land surface analysis scheme; this may explain ERA-Interim's relatively low MAE at DMI stations).

p.9, l.27: "variability in the downscaled RCM".

Done.

Figure 3: need to distinguish the two faint lines more clearly.

Done. (Note this figure has been moved to the supplementary material.)