The authors have presented a comparison of surface and near surface melt extent derived from passive microwave satellite sensors (SMMR and SSM/I) over the Greenland ice sheet and compared to temperature, melt water production, and snow liquid water content from two regional climate models (MAR and RACMO2) that incorporate snow/ice energy and mass balance models. As part of their methodology, the authors examined the best possible fixed brightness temperature (19H) threshold for use on the ice sheet by comparison with in situ GC-Net observations. Further, they examined the model output (i.e., temperature vs liquid water content) that provided the best comparison with the satellite data. They also examined biases in the models related to issues such as albedo parameterization and snow layer thickness.

Given the high correlation between the satellite and modeled results, the authors then extended the model back in time to demonstrate that recent years (2007, 2002, 1998) had surface melt that was unprecedented in the last half century. The article makes some important contributions. First, it provides a detailed discussion and assessment of the strengths and weaknesses of the various approaches that have been used to assess melt from passive microwave remote sensing, including issues related to the time of overpass. Secondly, the authors demonstrate that the regional climate model output can provide be used to extend the satellite record with some confidence. Finally, the authors highlight some important considerations regarding use of model output for comparison to the satellite record and of albedo parameterization within the regional models. Overall, I believe the manuscript makes an important contribution, although I do have some suggestions/questions detailed below.

General questions/concerns

1. The authors discuss strengths and weaknesses of the various microwave melt algorithms, but then merge a fixed TB threshold and the XPGR to create something called the ExtXPGR (3.3.2). It is not clear why this was done. The only explanation given is “Based on the analysis of the remote sensing algorithms, we evaluate the outputs of a new approach.”
2. While the 227.5K fixed 19H TB threshold does appear to provide a good fit to the GC-Net stations selected, it is not clear that these stations are representative of the entire ice sheet. (A map of the stations would be helpful.) Differences in accumulation rate, snow density, etc., may influence the TB threshold value. The authors should discuss how the use of a fixed threshold might influence the spatial comparison with the regional climate models. Perhaps this may be a reason for some of the regional differences between the models and the satellite-based estimates that were discussed in 4.2.
3. The authors indicate that the slope in the trend of surface melt from XPGR is not significant in coastal areas “likely” because of rainfall (4.1, and beginning of section 5). The finding that the XPGR melt trend is not significant in areas that might be expected to have greater rainfall and CLW is not sufficient

evidence to support this conclusion. Also, is the "XPGR based melt" in Fig. 9 referring to Abdalati's original definition of XPGR or Tedesco's impXPGR?

4. In 4.2, the authors state that "(t)he comparison with satellite-derived melt extent time series and those obtained with the two models suggests that at the beginning of the melt season, the snow albedo parameterization used in RACMO2 is not sensitive enough to wet snow conditions…" I do not understand how the passive microwave derived melt can tell us anything directly about the albedo parameterization. This statement needs additional explanation. Alternatively, other in situ and satellite sources of albedo data could be examined.

Minor questions/concerns

1. The authors should briefly mention other possible sources/methods of estimating surface melt from satellite, such as active microwave and thermal infrared.
2. At the end of 3.1, the authors indicate that the XPGR is sensitive to subsurface melt and the presence of liquid water in the snowpack with a refrozen surface. While the XPGR method may be more sensitive to subsurface melt, this statement could be applied to most of the microwave melt detection algorithms.
3. In 3.3.1, the two sentences beginning "But melt occurs only if the surface temperature…" are not clear.
4. The authors use the expression "melt threshold", but need to be more specific (e.g., the penultimate sentence of 3.3.1). They use multiple threshold values of brightness temperature and liquid water.
5. At the end of 3.3.2, the authors discuss the use of a 1.2% liquid water content threshold, rather than 1% used in previous studies. Can the model resolve LWC with sufficient accuracy to suggest that this is a meaningful difference?