Comment on tc-2021-384

Anonymous Referee #1

Referee comment on "A new Level 4 multi-sensor ice surface temperature product for the Greenland Ice Sheet" by Ioanna Karagali et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2021-384-RC1, 2022

This paper presents first a new data set of ice surface temperature (IST) merging different satellite products and assimilates afterwards this new IST data set to "improve" SMB estimations over the Greenland ice sheet. While the development of such L4 data sets is innovative and scientifically relevant as well as the assimilation into a SMB model, some improvements are need before potential publication in TC.

About the L4 data set:

- Line 122: Why a radius of 75km? What is the sensibility of using a larger/lower radius?

- Fig 3. Like fig4, all the time series should be shown on the same plot to better allow to compare the different data. Here, in view of the used -60°C +10°C vertical axis, it is impossible to evaluate if a product is warmer/colder than an other one. Moreover, nothing is said about the considered area. Is it the same one for each product or is it an integrated value over available area in each data set (which could be very different). A plot using at leat the same area for each day and each data set should be built. Finally, the differences in the passing time should be evaluated. The differences shown in Fig 2 could be due to the passing time which is different in each data set and therefore, these figures can not be compared for me.

- Figs 5-6: biases are systematically negative as LST (observed mostly during day) is compared to a daily mean (including night). Therefore, both product are not comparable for me because the passing time of LST is not representative of the daily amplitude of observed temperatures (This issue is moreover mentioned in the conclusion). The passing time should be considered to have a fair comparison or at least, only the day hours should

be used to compute the PROMICE average.

- Section 4.2: the mean LST from satellite products over 2012 should be compared in 2D by considering all the available data and by considering only the pixels present in each data set. The differences between the products are due to the passing time? the cloud mask which is different in each data set? or for the same area at the same time, observed LST is different? Moreover, the amount of missing data (still in 2D) in each data set should be compared instead of showing the number of aggregated observations in Fig 10.

About the assimilation:

- The passing time needs to be take into account for me. It is particularly relevant in May when melt occurs sometime only at local noon. By assimilating a daily average, this smooths the daily amplitude in the energy balance model and then the production of melt. It is particularly relevant at the beginning and at the end of the melting season.

- During the peak melting season, melt occurs every time and as IST is limited to 0°C, assimilating of not IST does not change the SMB results explaining why the focus has been made here over May.

- While the aim is to improve SMB and surface melt, nothing is said about the differences in the cumulated melt amount, runoff, SMB, … between the control and assimilation run. Only IST is compared between both simulations. But, in term of SMB and melt, it is not clear if this improves or not the results. As the temperature is forced in the SMB model, it is not very relevant for me to evaluate IST only. The melt extent should be compared with a microwave derived product for example to see the interest here of assimilating a daily product. Moreover, integrated over the whole season, what is the impact on the production of melt, runoff and refreezing?