Reviewer #1

We would like to thank the reviewer #1 for its constructive comments that will improve the paper. We have responded to all of them and will modify our paper accordingly. The majority of the highlighted issues of the paper come from its writing style. The paper will undergo a restructuring and will be sent to a native english speaker after rewording to reduce confusions. Our point-by-point answers (in bleu) follow as a supplement.

Major and general comments:

The paper generally identifies different sensitivities for the different assimilation parameters under consideration, however it is not attempted to identify a recommended assimilation algorithm. A basis for such an assessment could be to evaluate which experiment yields good agreement with the binary melt masks with minimal nudging. To this end it could be interesting to compare the values in Table 5 and Figure 11 to respective values diagnosed from the satellite-derived melt masks. Also Figures 12 and 13 might be extended and discussed in greater depth. And it would be interesting to map the total energy added and subtracted (separately) throughout the experiment within each grid box and to provide a budget for the whole domain.

> An attempt to identify the optimal parameters will be performed by analyzing the difference between the satellite products and the modeled number of melt days. For the energy balance, the exact value can not be inferred from the results as simulations need to be performed again for having these outputs. However, an estimation can be obtained by calculating the quantity of snow melted/refrozen compared to the reference simulation and multiplying it by the energy required to change snow/water state.

The experiments which use only one dataset should also be part of section 2 and should be discussed more systematically and in more detail. Due to the great number of experiments it is difficult to get an overview. It would be helpful to have one table with all experiments (table 2 does not include the one sensor experiments) and another one with some simple metrics such as the number of melt days and the total meltwater production over the whole melt season for each experiment.

> Table 2 will be extended to include one-sensor experiments. Another table will be added to give an overview of the results of the different experiments.

Liquid water in the snow pack is not necessarily indicating ongoing melting- it can also indicate past melt events with incomplete refreezing at night (e.g. for cloudy conditions or at greater depth)- this should be distinguished and also discussed with respect to the different duration of the melt season for the different satellite data sets. Also I would not use the term of "binary melt masks" - but something like "wet snow masks". Furthermore I think that *possibly valuable information is discarded when shallow penetrating data sets indicate no wet snow while the deep penetrating data set indicates wet snow. This is not necessarily a conflict but could occur after a melt period has ended but percolated water may remain liquid in deeper layers.*

> As satellites observe liquid water in the snowpack and not the melting phenomenon itself, using "*Binary melt masks*" as we did is indeed misleading and will therefore be changed to wet-snow masks. A similar change will also be performed for all the mention of a sensor "*detecting melt*". We discarded the satellite observations when shallow-penetration sensors do not detect liquid water but a deeper penetrating sensor does in a manner of consistency. This difference in detection can come from deeper water that is not observed but could also be caused by a difference in the sensitivity of the sensor or even a false detection. This is correct that valuable information is lost, but as we wanted to keep a balance between the complexity of the algorithm and computation speed, these cases are currently not taken into account.

As the satellites only detect presence of liquid water and not melt, I am also surprised that sensor penetration depth and assimilation depth are so closely linked here: I would not expect melt at depths of 1 m or more. I would rather have limited 1 the temperature nudging to a much shallower surface layer. However one could still compare water content (here I would use absolute and not relative values) down to the respective penetration depths of the available measurements and then trigger melt only in the surface layer. I wonder if there are reasons against such a strategy.

> The assimilation algorithm is based on Kittel et al. (2022) which nudges the snowpack temperature to the satellite penetration depth. As we do not have information on the depth of liquid water and its amount with the satellite imagery and the sensors are very sensitive to the presence of liquid water, the default warming depth is set equal to the penetration capabilities of the sensor. MAR can hold water in one layer until 5% of the air bubbles contained in the snow are filled with water before making it percolate to the underneath layer. By heating the snowpack uniformly, it is possible to have a more uniform presence of water.

Sometimes the word sensitivity seems to be used ambiguously. Most of the time it is used as in "simulated melt (whether more or less) depends strongly on the parameter choice" but for instance in l.14-15 it seems to rather mean "more melt is detected for a certain parameter choice"

> As this wording is confusing, we will revise it at each occurrence and change it accordingly in the revised version.

The words used for the assimilation parameters are unnecessarily diverse and confusing. I would recommend to consistently use something like melt water threshold and assimilation depth. The latter should not be named threshold in my view and it also should not be called penetration depth as this can be easily confused with the penetration depth of the individual sensors

> The name of the threshold will be uniformized throughout the paper and the name "*penetration depth*" will be changed to "*assimilation depth*" when it does not specifically refer to the penetration depth of the satellite.

The introduction is too unstructured, I give some specific comments below, but these should be only considered after sorting the different aspects in a linear fashion.

> The structure of the paper will be revised as well as the structure of the different parts of the paper. A strong revision by the co-authors and English native speakers will be performed.

Also the method part is hard to read and should be thoroughly revised

> As for the introduction, the methodology part will be revised to clarify and simplify to facilitate the understanding of this part. In this manner, multiple changes will be applied :

- Subsections describing each satellite dataset will be added.
- The description of the assimilation algorithm will be extended to follow Figure 5.
- The ll. 205-207 will be moved before the assimilation cases explanations.
- Subsections about the parameters will be corrected following the minor comments.

Maybe it is not a problem for people from the remote sensing community- but the paper is not easily readable for the wider community. For instance, datasets are sometimes referred to by their mission (Sentinel), the general measurement (radar, radiometer, scatterometer), some general classifications (active or passive sensors) or their instrument name (ASCAT)- this is unnecessarily confusing.

> It is true that in remote sensing, some dataset are referred to as the name of the mission or the platform rather than the name of the sensor (mainly when there is only one sensor aboard the platform, like Sentinel-1 or Sentinel-2). These specificities will be presented when the datasets are described, then datasets will more simply be referred to by the name of the satellite (i.e. "*Sentinel-1 (S1)*", "*AMSR2*" and "*ASCAT*") in the revised version to decrease the confusion it caused.

<u>Specific comments:</u>

Title: it is the MAR snowpack which is sensitive to the assimilation, not the satellite-derived surface melt. Maybe: Assimilation of satellite-derived surface melt into the regional climate *model MAR: sensitivity of the snowpack on the Antarctic peninsula to assimilation parameters.*

The title will be changed to: "*Sensitivity of the Regional Climate Model MAR's snowpack to the assimilation parametrization of satellite-derived surface wet snow on the Antarctic Peninsula*".

l. 1: please reword the whole sentence and possibly add 1 or 2 sentences. Here it would be good to introduce the problem (e.g. surface melt, runoff and accumulations cannot be directly observed on larger scales and models have uncertainties, remote sensing can only provide melt extent).

Abstract will be reworded and the problematics related to *in situ* observations and remote sensing will be included in the revised version.

l. 3: rather use "reduce uncertainties".

Noted and it will be changed accordingly in the revised version.

l. 18: maybe: second parameter mostly influences the duration of the melt period but it has only limited effect on the absolute melt water production.

This sentence will be reworded to: "*For the second threshold, the impact is more important on the number of melt days [days] rather than the melt production [Gt] itself*".

l.29: maybe: even moderate surface melt is thought to weaken ... leading to substantial mass loss.

Noted and it will be changed accordingly in the revised version.

ll. 31-34 too long, muddled, partly redundant.

This part will be removed in the revised version.

l. 35 climate models do not monitor (wrong verb), they do not comprise the ice body and only few include the snow pack.

Noted and it will be changed accordingly in the revised version.

l. 47: correct: induce -> induced.

Noted and it will be changed accordingly in the revised version.

l. 49: delete: In addition.

Noted and it will be deleted in the revised version.

ll. 54-61: this is better placed earlier in the section and merged with the earlier sentence about ice shelves.

The sentence will be included earlier as the structure of the introduction will be changed.

l. 62: be specific, here: melt -> melt areas.

Noted and it will be changed accordingly in the revised version.

l. 62: include Kittel et al. (2022) here and generally explain that the strategy is to warm or cool the snow pack in order to better match satellite derived melt maps.

Kittel et al., (2022a) will be added in the introduction and the general concept of its assimilation strategy will be briefly explained here.

ll. 66-72: this is too general and does not get to the point. I think you wanted to say that the different available products yield either poor spatial or temporal resolution and in contrast to Kittel et al. (2022) you test combining several products.

These lines are supposed to explain the main advantages/disadvantages of the passive/active sensor and that both can be used together to benefit from the advantages of the two types of sensors. As this is currently not clear, the part will be entirely rewritten.

l. 68: active / passive sensors should be explained in this journal before using these terms.

An explanation of passive/active sensors will be added at the same time as the subsection describing the remote sensing datasets in the revised version of our manuscript.

Section 2.1: this section could be better structured. Consider implementing subsections for each satellite/sensor type and providing a table with technical specifications (e.g. mission, sensor, resolution, revisit time, reference). It would be good to have a table with unique names for the four data sets and their technical specifications, and then to only use the data set names.

Subsections dedicated to each remote sensing datasets will be added as well as a table taking into account their specification to reduce the confusion induced by the name of the datasets.

l. 78: four data sets from three sources?

The AMSR2 data are splitted in two datasets : one with only ascending mode acquisitions and one with descending mode acquisition. *In fine* we have the

Sentinel-1 dataset, ASCAT dataset, AMSR2 ascending dataset, AMSR2 descending dataset.

l. 78: the fact that the strategy is to produce binary melt masks from satellite data and assimilate these should already be spelled out in the introduction and abstract.

It will be stated in the introduction and abstract that we produce **wet snow masks** we want to assimilate.

l. 90: this is confusing as here only three data sets are mentioned.

It will be clarified in the revised version as explained for the comments on line 78.

l. 94: "level 3 products" seems to be an unnecessary detail.

The level of a product in remote sensing stands for the operations that are already applied to the product before downloading it. Here, we specify the level as we did not calculate the temperature brightness myself. It also answers the question: "How does the author computed the temperature brightness?". Here we will simply specify the dataset we use by citing the location it can be retrieved.

l. 96: ascending and descending paths should be explained.

Ascending and descending will be explained before, in the AMSR2 subsection, following l. 78 comment.

eq. 10: how is TP measured?

We did not calculate TP as we use the provided level-3 product that contains the temperature brightness, as mentioned in l. 94 response to comment. Eq1 is there to remind us that the temperature brightness of a body will change if its emissivity changes but not its physical temperature.

l. 104: the reference to Fig. 3 is confusing here; I propose to refer to Fig. 3 in l.99 and refer to Fig. 2 at the end of l. 105.

Comments from anonymous reviewers #1 and #2 tend to suggest opposite directions (either simplify, or on the contrary go in more detail). Consequently, the reference to the figures may change. Considering the viewpoint of the second reviewer, having 3 figures (Fig. 2,Fig. 3, and Fig. 4) on melt detection, which is not the main subject of this paper, may be redundant. Fig 3 will be removed as well as its reference line 104.

l. 116: maybe better: sensors will indicate the presence of water .. by changes in the backscattering.

This sentence will be reworded to : “*It is possible to detect the presence of liquid water in the snowpack in Sentinel-1 images by identifying changes in backscattering coefficient σ0 through time (Figure 4)*”.

l. 125: this is not coming to the point: the -2.66 dB threshold is used in this study?

Yes the -2.66 dB is used on the normalized images of Sentinel-1 as threshold to detect wet snow.

p. 7, Fig. 3: use coastline contours also in the upper panels.

Coastline will be added in Figures that do not include them for better readability.

l. 140: explain or avoid the word scene in this context.

Noted, the word “*scene*” will be avoided.

l. 145: it is unclear to which part the word "else" is related to.

Sentences will be reworded . Sentences from ll.143-145 will be reworded to : “ *To create daily wet-snow masks, Sentinel-1 images of the same day were combined. In the case where three or more images overlap, the snow state is selected by a majority filtera and the acquisition time is defined as the mean time between the selected acquisitions. In the case where there are only two images and that contradict each other, the non-wet status is assumed. The acquisition time selected is then the acquisition time of the non-wet image*.”

p. 8, Fig 4: it could be interesting to see panel B after normalization.

It will be added in the revised version.

l. 146: A figure for the ASCAT data could be included, similar to Figs. 3 and 4. Also a reference for this data set is missing.

As explained previously, as the melt detection is not the main subject of the paper, adding another figure may complexify the paper although the ASCAT description will be extended. ASCAT dataset can be retrieved from the EUMETSAT data hub (https://data.eumetsat.int/data/map/EO:EUM:DAT:METOP:ASCSZF1B).

l. 157: "transfer between atmospheric part ... and the atmosphere" is this right?

It is a typo, as it should be “*transfer between atmosphere and soil*”.

l. *159: What is the typical vertical resolution in the upper 1.5m? Also please cover the percolation algorithm which seems to be crucial to understand the response in subsurface liquid water content.*

MAR is configured with a decreasing vertical resolution of the layers from the top to the bottom. The fists layers are typically at the centimeter size while under the first meter, we are at the meter resolution. The four first maximum layer thickness are respectively 2, 5, 10, 30 cm for example.

The percolation algorithm of MAR will be explained in more detail in the subsection 2.2 dedicated to model description. Here we worked with a maximum of the total layer weight being composed at 5% of liquid water before starting percolation.

Table 2 should be introduced in this section. Also it should include the single-data set experiments and I find experiment name MARa01−ku−02−c10 unfortunate as it does not indicate that here a different input is used and it does not indicate the assimilation depth of the third data set. Generally the experiment names are not very handy. I would suggest something like AsSdl for AMSR with shallow assimilation depth+S1 with deep assimilation depth and low water content threshold.

Table 2: Table 2 will be extended and placed in the introduction. The naming convention of the experiments have been changed and will be described in text to include the name of the sensor included for the assimilation. *E.g.* MARa01−ku02−c05 became $AsA_{01}S1_{05}AMA_{02}AMD_{02}$.

l. 173-174: check grammar.

Noted and changed accordingly in the revised version.

l. 175: correct: As up to three...

Noted. It will be corrected.

l. 183: is it possible to heat beyond 0 °C?

No, the snowpack temperature can not be heated beyond 0°C. Snow can not have a temperature higher than 0°C. At this temperature, snow is transformed into liquid water if more energy is available. Also, snowpack can not be cooled down under -7.5°C in our algorithm.

l. 187: either percolate into the ice or accumulate in the ice.

Noted and it will be changed to “accumulate in the ice”.

l. 190: better: discarded -> ignored.

Noted. It will be corrected in the revised version.

l. 192: shorter: if the two masks agree, the two observations…

Noted and it will be changed to "*If the two masks agree, the two observations are associated with the first case …*" in the revised version of the manuscript.

l. 199: more precise: at the same time -> within the same 3-hour time window.

Noted and iIt will be corrected accordingly in the revised version of the manuscript.

l. 206-207: maybe put this first.

Lines will be moved before the explanation of the different cases.

l. 213: check unit.

It should be Kg/m².

Section 3: Since the assimilation and the analysis are dealing with the snowpack it would be helpful to also evaluate precipitation and melt. Maybe compare to Wang et al. (2021).

As also pinpointed by the other reviewer, the MAR evaluation will be explained in further detail in the revised version and the dataset with which the evaluation is performed will be better stipulated. First, the observation data provided in Wang et al. (2021) does not cover the zone/studied period which makes the comparison complex. Second, the surface melt can also be evaluated by comparing the results with Jakobs et al. (2020). Results of this evaluation still need to be made carefully. Jakobs dataset also remains a modeled-based estimate, with its own biases and limitations, and therefore cannot be used as if there were in situ measurements, and so a reference for the model estimations. However, a short comparison of MAR_{ref} with the AWS used by Jakobs available on The Antarctic Peninsula can still be added to the paper.

Although, as we are testing the sensibility of the model, we only evaluate MAR without assimilation as we know that the value given after assimilation will differ from the observations.

As represented in the Figure here under (Figure R1.1) we compare the melt estimates from the AWS described in Jakobs et al. (2020). to the surface melt production of the 4 closest MAR pixels of the AWS. MAR has a tendency to overestimate some extremes of melting while simultaneously underestimating or overestimating the duration of periods during which the ice shelves are experiencing melting. Even if it is important to note that there can be a difference in altitude between the AWS and MAR pixels that explains the differences between the two datasets, this comparison also highlights the importance of nudging MAR to correspond to the remote sensing observation of wet snowpack.

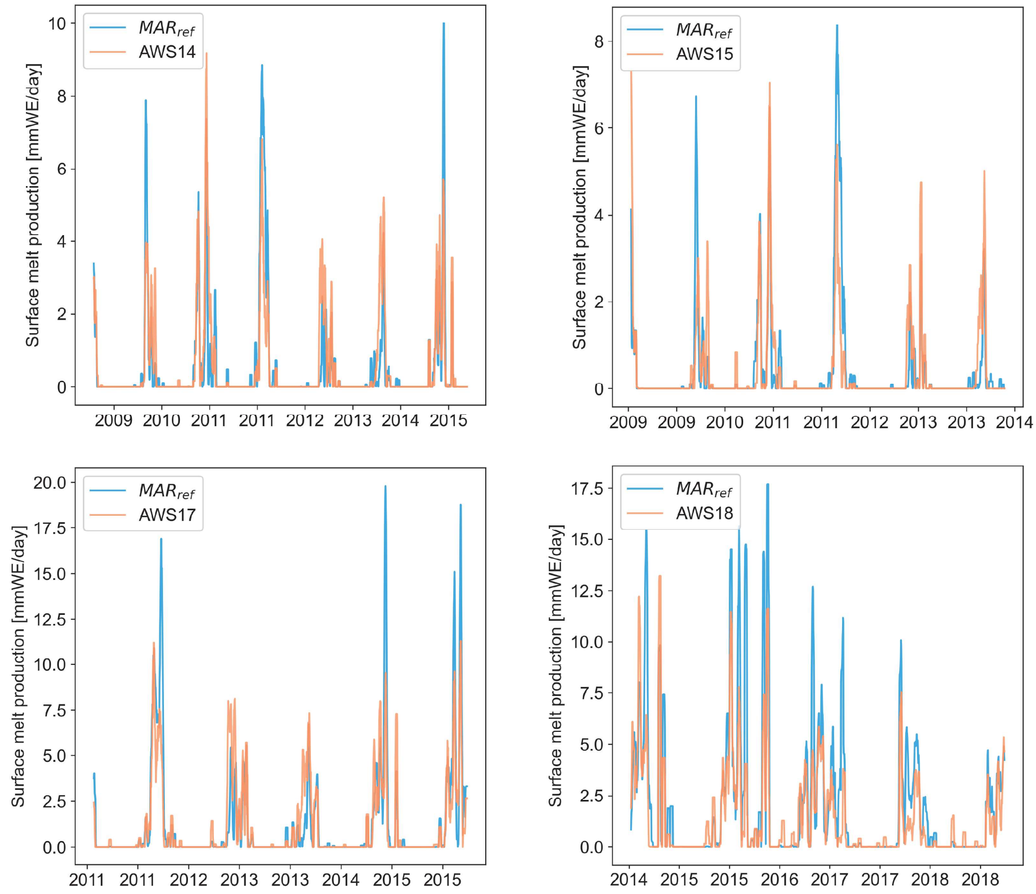

Figure R1.1 : Comparison of surface melt production as modeled by MAR without assimilation and estimated surface melt production from AWS 14, 15, 17, and 18 described in Jakobs et al. (2020).

l. 239: correlation (r).

Noted. It will be added in the revised version.

l. 249: better: a weak correlation and/or a strong negative bias.

Noted and it will be changed to "*A weaker correlation is observed in summer for long-wavelength downward radiations (r = 0.65).*" in the revised version.

l. 250: actually the bias is also strongest in summer (winter insolation should be weak anyway) and biases in net longwave radiation and net shortwave radiation almost cancel out and indicate underestimated cloud cover.

The biases are indeed caused by the cloud cover. The effect of clouds will be explained in more detail. The explanations of the effect of cloud cover in the simulation will be based on Kittel et al. (2022).

l. 255: "Combined with ...": unclear

It will be changed to: " *The implementation of cloud microphysics and the radiative scheme implemented (Kittel 2022b) suggest that MAR underestimates the liquid water path during summer when compared to Cloudsat-CALISPO estimates. Such underestimation is partially responsible for the LWD bias observed in summer.*"

Assimmean is an unfortunate name, as it suggests an experiment of its own right- I would propose Assim or mean(Assim). Also it should be stated here that three experiments were discarded in Assimmean

Name of the experiments will be changed in the revised version. $Assim_{mean}$ will be renamed to $\overline{Assim}$.

l. 258: correct is->are

Noted. It will be corrected in the revised version.

l.: 266: correct model-> simulation

Noted. It will be corrected in the revised version.

l. 269: gives different results from -> differs from

Noted. It will be corrected in the revised version.

Table 3: Evolution is not (but should be) mentioned in the caption- I understand that evolution is relative change due to assimilation in Assimmean, the name evolution is maybe misleading. It is not clear whether LWC and ρ are mean state or final states at the end of the period. Another column for LW C5m or some other deeper layer would be interesting. Also: replace mean value of the assimilations with mean value of the 16 assimilations selected for Assimmean

Table 3 will be changed to include mean value during summer and winter. Caption will be changed to : "*Change of surface melt production (ME), runoff (RU), surface mass balance (SMB), snowpack density (ρ), and snowpack liquid water content (LWC) for MAR_{ref} and the mean value of the assimilations ($\overline{Assim}$) over the Antarctic Peninsula. Variables are cumulated over summer (DJF) and winter (JJA) except for snowpack density which is the average density for the seasons.*

LWC and ρ are given for the 20 firsts centimeters and the firsts meter of the snowpack while the other variables are given as for the whole modeled snowpack."

As most of the variables remain constant or null during winter, the table discussion should not be different.

Figure 7: runoff could be shown in the same figure with a different y-axis on the right, maybe also highlight Assimref

Runoff and refreeze will be included in the figure.

l. 271: rephrase without the first part

"*Although there are divergences while using different parameters in the assimilation*," will be removed in the revised version.

l. 275: please check: 63.8 is the value for runoff according to table 3.

It should have been 66.7%. Noted. It will be corrected.

l. 279: the same -> almost the same

Noted. It will be changed to "*almost similar*".

l. 280 ff: clumsy, please rephrase.

l279 - 281 will be reworded to "Despite the fact that the relative increase in surface melt and runoff is almost similar for $\overline{Assim}$ and MAR_{ref} (66.7% and 63.8%, respectively), their absolute increase in Gt y-1 is not the same (+95 Gt y^{-1} and +21 Gt y^{-1}, respectively). This suggests that the snowpack can still absorb liquid water unless it reaches its maximum capacity."

l. 282: correct: depending on the energy balance

Noted. It will be corrected.

l. 284: densify -> densifies

Noted. It will be corrected.

l. 287: SMB is either snowfall + windrift + refreeze − melt − sublimation or precipitation + winddrift − runoff − sublimation also please clarify whether snow drift is represented in MAR.

In the case of this study, SMB= snowfall + rainfall - sublimation - runoff as the blowing snow module was not active for these simulations.

Blowing snow module was turned off to increase simulation speed and to not impact the sensitivity.

l. 292: please specify "deeper".

The word "*deeper*" will be removed as it brings nothing to the discussion.

l. 293: not all ice shelves exhibit lower liquid water quantity.

l. 293 will be reworded to " *With a denser snowpack, firn air content is reduced and there is less space for liquid water to be absorbed. Therefore, despite the increase in surface melt production, the assimilation process may lead to a decrease in the amount of liquid water retained in the snowpack. This is because the assimilation causes a reduction in the snowpack's capacity to retain water*."

l.295ff: Table 4 should be discussed in more detail: there is no explanation given, why LWC for Larsen C is increasing. Also for individual ice shelves it is not true that relative changes in melt and runoff are of similar size. It is particularly not true for Wilkins where additional melt almost entirely becomes additional runoff. For a better understanding it could help to map the degree of saturation in the upper snowpack and to look at different stages of the melt season. A deeper interpretation of Table 3 and 4 is also difficult due to ambiguous variable definitions (see below)

l 295 will be reworded to : "*All three highlighted ice shelves (Larsen C, Wilkins, and Georges VI) are expericening a increase in surface melt production and runoff (Table 4). Similar conclusion drawn for the AP can also be applied for them.*"

The evolution of LWC will be studied in more detail as some bias induced by the nudging of the snowpack temperature could be overlooked by only using this table. The data presented in Figure R1.2 indicates that the Larsen C snow pack has a greater amount of liquid water following assimilations. This can be attributed to the fact that the snow pack accumulates water at an earlier stage before the second peak of melting and retains it for a longer period. It should be noted that the values in Figure R1.2 are one order of magnitude higher than those in Table 4 as the quantity of liquid water has been summed over the layers up to a depth of 1 meter, rather than being averaged.

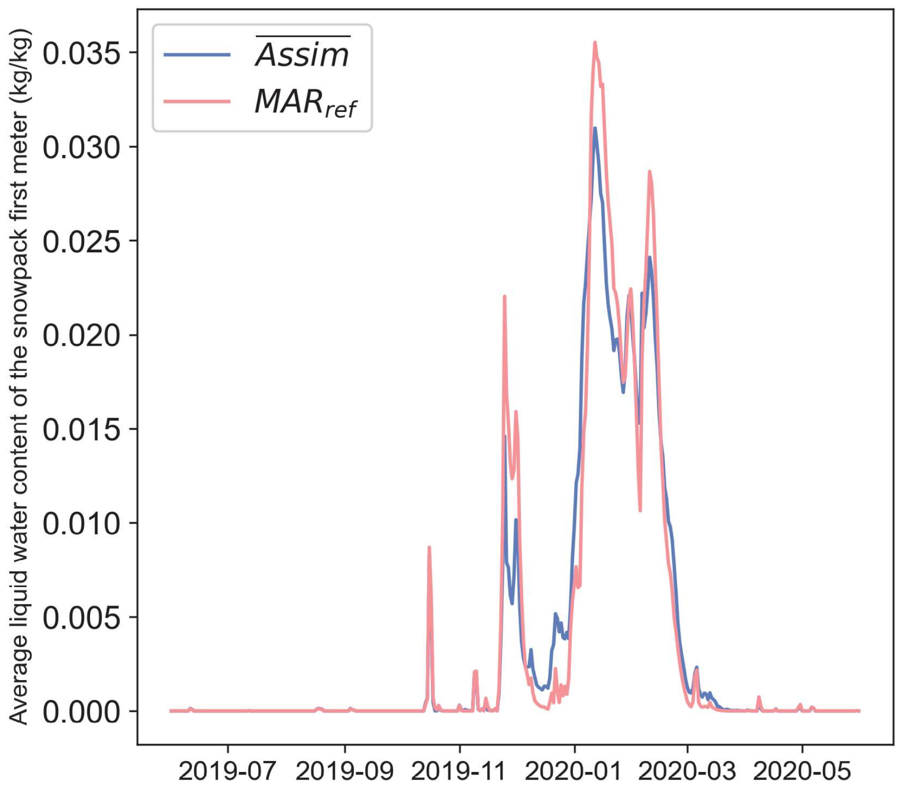

Figure R1.2 : Comparison of the average cumulated liquid water content of the first meter of the Larsen C snowpack as modeled by MAR without assimilation and the average value of the assimilations.

ll. 299-302: this is completely unclear to me.

It will be reworded to "*Except for the liquid water content, the snow-related variables (ME, RU, SMB and snowpack density) of the model have undergone significant changes, causing MARref to fall outside the range of the various assimilations. As a result of increased surface melt production, there is an increase in runoff and a subsequent decrease in surface mass balance. This increase in runoff is attributed to the compaction of the upper layers of the snowpack, which reduces its capacity to absorb meltwater.*".

l. 303: It is not really surprising that the mean of the assimilation experiments is close to the central reference experiment. However without evaluation this is not necessarily meaning that this is more realistic than other members.

In this paper, we are not trying to obtain better melt estimates with the assimilations. The main purpose is to test the sensitivity of MAR to the parameters of the assimilation in the aim of evaluating uncertainties on the simulated melt amount. Parts of the articles will be rewritten to state it more clearly. In addition, we want to observe where/when the model is different from the binary mask created with the satellite observations and what are the impacts to match this liquid water extent in

MAR. This point will be discussed in more details and will be more explicit. Here that statement will be reworded to "*In the end, the results illustrate that $Assim_{ref}$ is the closest simulation to $\overline{Assim}$, and makes it an appropriate candidate when computational resources are limited and only the effect of the assimilation want to be studied, not the sensitivity of the model*."

Table 4: The caption to this table is sloppily formulated. As in table 3 not all lines and columns are well defined or self explanatory and there is room for guesswork but no sound basis for interpretation. Also I wonder why LWC is consistently one order of magnitude smaller than in Table 3.

Caption of Table 4 will be rewritten following the new caption of Table 3 and σ will be removed from the table.

In Table 3, we present cumulated values over the whole Peninsula while in Table 4 we present cumulated values only over individual ice shelves. Thus, if we add the values of LWC over the ice shelves and the rest of the studied zone, we will have the values of Table 3. This information will appear clearer in the revised version.

l. 310: refreezing is indeed releasing energy and heating the ambient snow!

Yes, refreezing releases energy and heats the snowpack but the quantity of heat released is not sufficient enough to heat up the layers under one meter as even with higer refreeze, $Assif_{ref}$ snowpack temperature under one meter remains similar to MAR_{ref} snowpack temperature. The colder layers will eventually cool down the snowpack.

Please revise ll. 310-314.

This part will be reworded to: "*First, the available energy in the system is consumed by melting processes, preventing the layer under 1m from heating up. A colder snowpack constantly needs larger nudging to reach the melt threshold. The second point is that due to the lower saturation of water in the lower layers, the upper layers become saturated with less water because of densification during melt events, resulting in increased runoff and faster percolation of the water into deeper layers, outside of the assimilation depth range. If the model were to retain liquid water in its top snow layers for a longer duration, it would require less nudging to match the RS datasets*"

l. 310: not sure what prevails means here. Maybe prevail->prevent?

This sentence will be removed in the revised version as presented in the comment above.

l. 317: please specify what exactly qualifies results to be improbable. Such an exclusion criterion should be defined beforehand. And the exclusion of the members should be stressed when introducing table 2 in the method section.

Exclusion of the simulations will be introduced near line 264 when Table 2 is discussed.

Here, the exclusion criterion was that the Antarctic Peninsula was experiencing negative SMB one order greater than the other simulations when cumulated over the whole melt period on the Antarctic Peninsula. This behavior is not observed in datasets used in other studies (Kittel, 2021; Chuter et al., 2022) and thus may introduce a bias in the comparisons.

l. 333: lesser -> less or smaller.

It will be corrected in the revised version.

Table 5: Which are the experiments considered here? Are these numbers the same for all experiments with α = 0.1 and α = 0.2 ? Also maybe noteworthy: number of melt days larger for MARref than one of the assimilations on Wilkins.

All these experiments are considered. It is the mean values for the assimilation with α = X that is shown. Caption will be changed to "*Comparison between the melt season length and number of melt days modeled for the three studied ice shelves for MAR_{ref} and the average number for assimilations depending on their α for the 2019-2020 melt season.*" to make the message clearer.

l. 343: is this referring to the whole 20m snow pack?

Indeed this important information is missing. Only the first meter of the snowpack is considered, and not the full modeled snowpack.

l. 347: I don't find these numbers in Table 5- I stop reading this paragraph here.

This will be corrected in the revised version as the value given refers to Figure 11. L347-348 will be reworded "*By computing the mean value of each pixel number of melt days of the ice shelves, it was found that the largest deviation occurs on Larsen C, with an increase of 15 melt days. The other two ice shelves exhibit comparatively smaller differences, with Wilkins and Georges VI experiencing an increase of 8 and 9 melt days, respectively.*"

l. 364: this cycle -> daily melt - refreezing cycle?

Yes, the daily melt-refreeze cycle.

l. 373: is other frequency here higher frequency?

In this case, the "*other frequency*" should have been "*another frequency than AMSR2*" as it is referring to ASCAT. With a better description of the sensors, there will be less confusion. ASCAT frequency (~5GHz) is lower than AMSR2 (~19GHz).

Section 5: this probably needs to be rewritten anyway after the other parts of the manuscript have been revised.

With the new insight on the paper, Section 5 will be rewritten in the revised version. Nonetheless, the conclusion will remain mainly the same. Changes that will be included in section 5 are a more clearer take-home message, a conclusion on the change in liquid water content induced by the assimilation and a small discussion about the evaluation of current surface melt production of the model without assimilation.

l. 394: effect of assimilation was not studied here.

"*However, the assimilation of surface melt occurrence has a small impact on the atmosphere*." will be removed in the revised version.

References:

Chuter, S. J., Zammit-Mangion, A., Rougier, J., Dawson, G., & Bamber, J. L. (2022). Mass evolution of the Antarctic Peninsula over the last 2 decades from a joint Bayesian inversion. *The Cryosphere*, *16*(4), 1349–1367. https://doi.org/10.5194/tc-16-1349-2022

Jakobs, C. L., Reijmer, C. H., Smeets, C. J. P. P., Trusel, L. D., Van De Berg, W. J., Van Den Broeke, M. R., and Van Wessem, J. M. (2020). A benchmark dataset of in situ Antarctic surface melt rates and energy balance. Journal of Glaciology, 66(256), 291–302. https://doi.org/10.1017/jog.2020.6

Kittel, Christoph. (2021). Kittel et al. (2021), The Cryosphere : MAR and ESMs data [Data set]. Zenodo. https://doi.org/10.5281/zenodo.4459259

Kittel, C., Fettweis, X., Picard, G., and Gourmelen, N. (2022a). Assimilation of satellite-derived melt extent increases melt simulated by MAR over the Amundsen sector (West Antarctica), Bulletin de la Société Géographique de Liège, 78, 87–99, 2022.

Kittel, C., Amory, C., Hofer, S., Agosta, C., Jourdain, N. C., Gilbert, E., . . ., Fettweis, X. (2022b). Clouds drive differences in future surface melt over the Antarctic ice shelves. The Cryosphere, 16 (7), 2655–2669. https://doi.org/10.5194/tc-16-2655-2022