Author responses to Referee #2

Major Comment 1 (MC1) "My main comment concerns the dating of reflection layers. Here, the dating relies only on big assumptions made on stratigraphy characteristics. This technique is based on differences in winter and summer snow due to changes in atmospheric conditions and radiative fluxes. I suppose that dating of reflection layers and NP data is accurate but there is no comparison with stake networks, or with a clear "absolute" dating based on anthropogenic radionuclides or volcanic horizons at several cores. Since layers sometimes mismatch at several intersections, or are excluded (around iSTAR sites 2 and 19 for instance), the final dating may be not fully robust. I understand that the authors define a layer age uncertainty of +-1.4 year and assess the associated uncertainty in the surface accumulation, but is it possible that the layer dating uncertainty exceeds 1 year? In particular, melt or rain is expected to occur mainly in summer, but is it possible that significant surface melt (or rain) occurrences occurred at the beginning and end of summer but were separated by an "extreme" solid precipitation event (Turner et al., 2019), leading to a 2 maxima in density and in other snow characteristics used for the layer counting? Is there any snow erosion, which could remove the surface layer at locations in the area? Are there any stake farms or well constrained ice cores (with an absolute age of one or various layers) in the area, on which the authors could validate their estimates in snow accumulation?"

We agree that the concerns in MC1 about the interpretation of measured density peaks and internal reflection layers as annual markers are justified.

We propose to address the concerns in MC1 in the methods sections for the NP and ASIRAS data as follows:

Formation of melt layers can be produced at PIG in response to melt events. We shall include a reference to Scott et al. (2010), who observed an exceptional ice layer at lower PIG at 20 m depth. However, no other ice layers (> 1mm thick) were observed at their discussed core site. We shall also include an additional reference to Nicolas et al. (2017) to address the potential impact of large warm air intrusion events to the West Antarctic ice sheet on surface melt. In this connection, we expect that large scale surface melt episodes at PIG are connected to the specific coupling of atmospheric modes and therefore do not follow a frequent seasonal pattern. We cannot rule out, though, that seasonal surface melt events are more common near the coastline and therefore may result in the formation of ambiguous annual peaks in the NP and ASIRAS profiles, which we shall include in the text of the method sections (see also our response to MC4 of Referee #1).

Morris et al. (2017,Sect. 4) provide a detailed discussion on the applied methods for interpretation of annual density markers from the iSTAR NP measurements, which we shall refer to in the NP methods section of this study. We shall also refer to Morris, 2008 with regard to the applied calibration equation of NP measurement, which is based on theoretical considerations and validated using laboratory experiments and gravimetric data from ice cores. This calibration was performed independently from the iSTAR cores.

Surface snow erosion or missing formation of annual markers cannot be ruled out for the NP data according to Morris et al. (2017) and for the ASIRAS data. For the latter, we are pointing out on Line 128 that bridging of data becomes necessary where annual layers cannot be traced, e.g. due to wind erosion. This may also result in initial mismatches between the traced layer at some intersection, which we attempted to correct by retracing adjacent layers (see our reply to MC1 of Referee #1). Therefore, it is important to outline that we aimed at producing a consistent layer trace which matches at all intersections. We shall make this point more clear by using the term "correcting" instead of adjusting in the text, as proposed in our reply to Referee #1.

In terms of the robustness of our layer dating approach, we benefit from the intercomparison of different annual markers (i.e. reflectivity, density, and H_2O_2 at core locations) at iSTAR sites, which obey different physical processes in their formation. The photochemical H_2O_2 tracer (Line 81) follows a clear annual cycle in the PIG cores, which we can add to the text. Wind erosion may remove peaks in the H_2O_2 profile. This is less of a concern where surface accumulation is high. Locations exposed to pronounced wind erosion are visible in the observed stratigraphy from the ASIRAS soundings (Line 131), but annual reflection layers remain distinguishable at iSTAR locations, which supports the assumption of undisturbed formation of annual H_2O_2 markers here. According to internal communication with the iSTAR drilling team, no absolute volcanic reference horizon was accessible for the processed iSTAR cores. Stake farms do not exist for the considered region, but single stake measurement were performed between T1 and T2.

With regard to the assumptions made on stratigraphic characteristics of the observed internal reflection layers, we propose to refer to additional earlier studies in the introduction section (Line 30): The work of Arcone et al. (2004 and 2005) provides theoretical and observational evidence of the formation of continuous isochronal reflection layers throughout West Antarctica.

The question "is it possible that the layer dating uncertainty exceeds 1 year?" should have already been addressed in Sect. 2.4. It is true that we cannot completely rule out the possibility of systematic errors in the layer dating which may result from ambiguous annual signals or unintended layer changes in the manual tracing, but according to our response to MC4 of Referee #1 we shall clarify that an associated layer dating uncertainty of +/-1 years has already been taken into account for the estimated Δt = 1.4 years uncertainty of the layer age. It is highly unlikely that such systematic error persists for all data samples (i.e. NP, core measurements, and ASIRAS) at all sites which we use for the estimation of Δt and so we consider the error estimation above to be reasonable in terms of the robustness of the layer dating approach of this study.

MC2 "Neutron probe is a really interesting way to retrieve snow density and the authors clearly took profit from this technique in the past and in the present paper. However, data rely on a few calibration steps. I had a look to previous papers from Konrad et al., 2019 and Morris et al., 2017, and I did not really understand how density data were validated before being used in the present paper. My concern is because in Figure 2 we observe that snow density is 550 kg m-3 at 7m below the ground level, whereas it is 600 kg m-3 at the same depth in Morris et al., 2017. Snow density in firn cores from Konrad et al., 2019 are hard compare here because their Figure 3 is developed until 50 m. Could the author describe whether they calibrated the NP snow density data with snow pit data in the present paper or not? If not, is there a difference in the density/depth relationships between (Konrad et al., 2019), (Morris et al., 2017) and present paper. If snow density from NP measurements is biased, how will this impact the final SMB values?"

We are afraid that different units for depth, i.e. water equivalent depth (Fig. 2, Morris et al. 2017) and geometric depth (Fig. 2, this study) led to some confusion. Conversion of geometric to water equivalent depth yields an approximate density of 600 kg m^-3 at site 21 and 7 m level for this study. Even though, some of the referenced earlier studies use water equivalent scales, we prefer to keep SI units in this study to avoid potential pitfalls from the conversion to a specific reference scale as these are unambiguous. Nonetheless, the difference is well spotted by the referee and highlights the required caution to be taken when comparing different scales.

MC3 “Concerning the kriging method, is it worth using northing, easting, and elevation as explanatory variables? Would it be more relevant to use the distance from the coastline as explanatory variable? Or even the distance from Amundsen sea coast?”

Our initial motivation was to follow the methodology by M14 as close as possible for our data comparison. It is an interesting idea, if changes to the parametrisation of the regression model improved the statistical characteristics of residual SMB to the regression surface in our case. However, limiting the explanatory variables to the coastal distance would also limit the degrees of freedom in the regression model. We can only speculate at this point, but orographic effects in response to atmospheric circulation patterns may not be adequately captured by such a parametrisation.

Another potential explanatory variable may be temperature (Arthern et al., 2006), which affects saturation vapor pressure, which again correlates with accumulation rates. In turn, any correlation between elevation as explanatory variable in M14 and temperature makes both variables interchangeable for the regression model.

However, due to the persisting skewness of the residual SMB values based on our regression (Line 197), we completely rejected this approach for our analysis and used a logarithmic data transformation of the SMB samples instead (Line 202). It seems that this point is not clearly pointed out in our description and will be added accordingly.

MC4 **“C**omparison of ASIRAS data with RACMO2.3p2 simulations are really interesting but differences are not fully justified/explained in the text. Since differences are model dependant, it would be interesting to see potential differences with another model used in Antarctica (e.g., COSMO-CLM, HIRHAM5, MAR3.10, MetUM , see Mottram et al., 2020). Since Agosta et al. (2019) proposed potential justifications of the differences existing between RACMO2.3p2 and at least the MAR model, I believe that a comparison with the MAR model would make sense here. Indeed, in Agosta et al., 2019, large differences between RACMO and MAR are observed in regions where the RACMO2.3p2 model presents the largest differences with the ASIRAS data. A quick comparison could be relevant to discriminate whether the precipitation formation, advection of hydrometeors, and sublimation of precipitating hydrometeors (Agosta et al., 2019) are important or not in the PIG area. Data from Agosta et al. 2019 are available here: https://zenodo.org/record/2548848#.X0St8TXgphE If the authors are interested in higher resolution simulations, data from a more recent paper from Donat-Magnin et al., 2020, focusing in the Amundsen region are also available at: https://doi.org/10.5281/zenodo.2815907.”

Even though this study uses model data to augment missing observations, its main emphasis is on the observations and applied methodologies to provide a means for validation to the model community. While MAR may perform better in some regards, RACMO may do so in others. From the standpoint of an observational study this is difficult to judge and already exceeds the scope of this study. We are also concerned that the inclusion of a model-to-model comparison will impact its conciseness (see also MC6). However, if the MAR model helps reduce the elevation depended offset to our observations, we may include this point in a revised version of Figure 9 and refer to the findings above.

MC5 “The paper largely describes differences with Medley et al. (2014) paper, but the authors never include any figure presenting the differences. I propose to include a map presenting ME14 route, SMB results and differences with the ASIRAS data.”

We are afraid that the inclusion of further maps showing the M14 SMB results and differences with the ASIRAS data would make the paper less concise and may not contribute to its readability as concerned in MC6. Any side-by-side comparison between both maps is still possible based on the published results in M14 and the main focus of our intercomparison is on the estimated total mass input between both studies, which is considered in the provided tables of this study.

However, we think that it is useful to include the flight tracks of M14 in Figure 1 to aid any side-by-side comparison of the regional SMB distribution between both studies. If access to the flight coordinates is granted by the authors of M14, we include the flight tracks in the revised manuscript version.

MC6 "The paper is sometimes quite hard to follow for a non-expert of this area. Different datasets are used here, and the difference between ISTAR/ASIRAS and ME14 is not always clear. The authors use a many acronyms. I suggest that the authors include a table where they clearly describe the difference between ASIRAS/iSTAR data used here, and more particularly the difference with the ME14. For instance, Table 1 presents different radars (ASIRAS, pulseEKKO PRO GPR , CReSIS Accu-R) part of the information is given in the caption, but perhaps the authors could also precise in the table if field campaigns were deployed on the ground or by plane, over what distance? which were the reflection layers used for SMB estimates? which density measurements were considered (NP? Firn cores?)? how was performed the dating of the reflection layers?"

We appreciate the suggested improvements on the readability of the manuscript and further clarification of considered data products and we consider these suggestions for the revision.

MC7 "Figures could display the route followed by ME14, and where GPR from Konrad et al., 2019 was carried out."

As suggested in our response to MC5 we include the flight track followed by M14 in Figure 1 of the revised version, if data access is possible.

The ASIRAS flight track follows the iSTAR traverse, including the GPR measurements (Fig. 4 in Konrad et al., 2019), in most parts. We therefore prefer to limit the displayed routes to datasets, which we compare in this study to maintain readability of included maps.

Minor Comment 1 (mC1) "Thus there is no evidence of a secular trend in mass input to the PIG basin." => please be more accurate because secular may be misinterpreted here. I suggest to replace secular (here and elsewhere) by decadal."

We agree that any statement about secular trends depends on the considered timescale and needs clarification in the text. It is our intention to express that we could not find evidence of a pronounced shift between both studies which may be of non-periodic nature in response to the steady acceleration of dynamic ice loss of PIG since the 1970s. We think that replacing the term secular with "decadal" would be more limiting in this regard, therefore we prefer to clarify our intention for using this term in the text (Line 352).

mC2 "in particularly"

correct spelling, "in particular~~ly~~"

mC3 “Lines 38-41: firn cores are only used to retrieve the depth of dated snow layer? Are they used to calibrate the neutron probe density data?”

As mentioned in our reply to MC1, firn cores were not considered for the calibration of neutron probe measurements. However, firn cores were used to guide interpretation of the records at two sites.

mC4 “Figure 1. ASIRAS-iSTAR survey : please also include the location of ground GPR observations from Konrad et al., 2019?”

See our response to MC7

mC5 **“**Line 68, the authors write: ” Due to the reported consistency between the GPR and airborne SMB measurements in Konrad et al. (2019), we limit the comparison of our results to the basin wide estimates by ME14”. I feel that a figure (perhaps in the supplementary material) showing the different radargrams could help the reader. A quick data comparison, before interpolation, could also be done to see how snow density and radar data uncertainty impact the final SMB value.”

We agree that a comparison of radargrams from different platforms can be useful for an ongoing discussion on the error assumptions made by different studies. However, we are cautious about the expectation that this may be done by means of a quick data comparison as suggested in mC5. The radar studies of M14 and Konrad et al. (2019) use a different reflection layer for their analysis compared to this study, which already impacts the uncertainty of their SMB estimates. Local noise in the stratigraphy varies along the flight track and impacts related uncertainties in addition. Our error discussion builds upon the developed methodologies in M14 and Konrad et al. (2019), and already considers impacts from the uncertainty in the assumed regional density profile and dating precision for the selected reflection horizons. Figure 10 (a) already gives a benchmark for other studies to the estimated measurement uncertainty of our results. However, we think that it is of help for potential future studies to add the ASIRAS radargrams, interpolated, and non-interpolated reflection layer traces in units of annual SMB and TWT to the PANGAEA repository.

mC6 “Morris et al. (2017) applied an automatic annual layer identification routine to their snow density profiles and used the annual H_2O_2 peak depths as an additional guidance for the annual layer dating.” => is it possible to observe the removal of on year of snow in the PIG area (due to erosion) or multiple summer maxima?”

See our response to MC1, where we address the possibility of wind removal of annual markers and multiple summer maxima in response to melt events.

mC7 “Here I understand that the authors did not consider the density obtained from the firn cores to compute the final SMB. What is the difference in the final SMB if the authors use the density from firn cores to calibrate the snow density profiles?”

We confirm that we only use density profiles from NP measurements for the estimation of the regional density depth relation (see our response to MC1). We consider that the NP calibration has already been thoroughly elaborated in Morris, 2008 (see our response to MC1). Indeed, iSTAR NP profiles were compared with the core profiles to check that there was no systematic difference between them. Hence, we expect that the impact of reducing the number of sites will be larger to the fit of the regional density profile than the difference between the collocated NP and core measurements.

mC8 "Line 89: 43 profiles => I suppose this means that profiles were done twice at the 22 sites?"

All NP density profiles are composed by two adjacent profiles per site, with the exception of site 2 (Line 78). We agree that counting the adjacent profiles separately may confuse the reader (i.e. 2x21 adjacent profiles plus the single profile at site 2) and we shall only count the resulting concatenated profiles in the revised text.

mC9 "Line 92: is there any relationship between snow density profile and Accumulation/ Temperature as suggested by Herron & Langway 1980 equation?"

There is a relation between snow density profiles and accumulation and temperature, which is fully discussed in Morris et al. (2017). While the Herron & Langway 1980 (H&L) equation gives good predictions for stage 1 and stage 2 densification at the iSTAR sites, large departures to the H&L equation are observed for the transition zone between stage 1 and 2 and requires the consideration of a new model equation as described in Morris et al. (2017).

mC10 "Line 120: how does hoar produce thin ice layers? Is it possible to have short rain our melt events in spring or late summer? Could wind erosion or sublimation create wind crusts in this region?"

The first question "how does hoar produce thin ice layers" should already be covered by the provided reference to Arcone et al. (2004) on Line 120, i.e."[...] a hoar layer, which frequently occurs beneath thin ice layers because it is the source of vapor that creates the ice layer"

See our response to MC1 with regard to the possibility of rain or melt events at PIG.
With regard to the final question, "Could wind erosion or sublimation create wind crusts in this region?", wind crusts are more often observed in region with lower accumulation, e.g. on the polar plateau, than in a coastal area like PIG, where synoptic precipitation prevails.

mC11 "We assign an estimated date to the mean value of N points" => do you mean "the mean value of depth"?

Well spotted, the word "depth" is missing and will be added in the revised text.

mC12 "Line 198: "may by due"

Thanks, will replace "by" with "be"

mC13 "Line 275: please discuss this sentence according to Agosta et al. (2019) results. Indeed, according to this publication, sublimation of precipitating hydrometeors are missed at low elevation in RACMO2.3p2. This could justify that RACMO 2.3p2 presents a positive bias at low elevation. Conversely, they suggest that MAR snowfall rates generally exceed those simulated by RACMO2.3p2, by more than 30% on the lee side of the West AIS (Marie Byrd Land toward Ross ice shelf), and close to crests at the ice sheet margins. Here, a comparison with MAR could be interesting."

We agree that our discussion benefits from a reference to the findings of Agosta et al. (2019) and will be included in the revised text. To keep the scope of this study concise, we may prefer to limit our analysis to the RACMO based simulations as discussed in our response to MC4.

mC14 Lines 288: The authors refer to the ASIRAS or the Hydrid SMB estimates, but line 296 they refer to the ASIRAS vs. the ME14 one, whereas they refer to the Hybrid estimate at line 298. The difference between these estimates is not clear. Why do the authors use ASIRAS / hydrid estimates in different parts of the text?

Line 296 should refer to hybrid estimates as well. We will add the word "hybrid" accordingly.

mC15 "Line 325: "but also measured density and strainrate profiles suggest a mean annual SMB of 200 kgm−2yr−1 based on theoretical grounds, which both are in a better agreement with the collocated ASIRAS based results." => Is it possible to explain briefly the way this value of 200kg m-2 is computed?"

The value of 200 kg m^-2 was computed from the Herron and Langway Stage 1 equation based on the measured density and strainrate profile. We can address this in the revised text.

mC16 "Line 328- 331: please give elevation of sites 2, 18, 19. It would be interesting to discuss the differences between the RACMO2.3p2 and MAR models here (is there any potential explanation related to precipitation formation /advection /sublimation of precipitating particles to the ground described in both models)."

We can add elevations for sites 2, 18, and 19. While the suggested model comparison sounds exciting, it appears to us better suited for a separate model intercomparison study, which may also include high resolution SMB measurements for comparison. (see also our replies to MC4 and mC13)

mC17 "Line 352: "secular" => decadal?"

See our reply to mC1**.**

mC18 "Inspection of the artificial cluster highlighted in Fig. 5 revealed" => replace by fig. 8?"

Well spotted, it should have been Figure 8

mC19 "Table 2: please include site coordinates and elevation."

Due to the redundancy of site numbers in Table 2, the inclusion of site coordinates and elevation would yield to considerable white space. We suppose to refer to Table 1 in Morris et al. (2017) in the caption instead. All requested information is presented here.

mC20 "Table 3: Are the basins from Mouginot et al. (2017) similar to those given by Rignot et al. (2019)? If not, perhaps it would make sense to use the most recent basins."
Yes, they are. See Rignot et al. (2019): "The […] drainage basis are available at National Snow and Ice Data Center (NSIDC), Boulder, CO as MEaSURES-2 products". The Mouginot et al. (2017) reference corresponds to the requested citation for using the data provided by NSIDC.

mC21 "Table 5: do values refer to results written as OLK(NN) in Figure 7?" and
mC23 "Figure 7: Please define NN"

No, OLK(NN) refers to the PP-plot by comparing *observations* against the nearest neighbour estimates to the *observations*. Ideally, probability density functions should agree between both of them, which becomes evident by the 1:1 match in the PP-plot. However, if we only consider nearest neighbour *estimates* for the ASIRAS measurements, we would still miss a large fraction of the PIG catchment. This is why we have to expand our estimates to points beyond nearest neighbours of *observation*. By extending the allowed range of *estimates* to the *observations,* which they are based from, the PP-plot should still remain a straight line, which is the case for the applied ordinary lognormal kriging estimation. The indicated 100 km and 190 km ranges indicate the transition zone from estimates based on observations only and estimates based on RACMO.

Indeed, we missed to define NN as nearest neighbours in the text. This should also clarify the role of varying sample populations for Figure 7.

mC22 "Figure 5: please include ISTAR site numbers on a) and b) to help the reader in retrieving the locations cited in the text."

We can do this for the revised version.

mC24 "Figure 8: I suggest to include the elevation contours."

We initially tested elevation contours on top of Fig. 8 but we had the impression that it affected the readability of the images. We may, however, test larger increments between contour lines in a revised version of Figure 8.

"Is it correct to write Eastings when the x-axis is related to the north-south direction, and Northings for the y-axis when it is on the east/west direction?"

Eastings and northings are a common convention for geographic Cartesian coordinates and can be found for this region with the same orientation in the literature (e.g. Figure 2 in Grima et al., 2014). However, we agree that the northing/easting convention can be confusing when compared to actual directions. Another possibility, which we can do is to replace the axes annotations with "polar stereographic x [km]" and "polar stereographic y [km]" to avoid any confusion about directions.

"Is it because northing, easting, and elevation are used as explanatory variables? If yes, is it not more relevant to use the distance from the coastline as explanatory variable? Or even the distance from Amundsen sea coast?"

We spot a remaining misunderstanding in this question, i.e. "[…] is it not more relevant to use the distance from the coastline as explanatory variable?" In contrast to ME14, we do not use a regression model to produce residual SMBs for the kriging method (see our reply to MC3).

Indeed, straight lines are indicative of artifacts in the interpolation scheme, which limits the estimation to a fixed number of nearest neighbours as suggested by Yamamoto, 2007.
Again, these neighbours are split up according to the quadrant criterion, which can produce such artifacts depending on the number of nearest neighbours. We tested increasing the number of nearest neighbours, which reduces these artifacts but yield to a departure between *observations* and *estimations* in the PP-plot, which we use as a benchmark for the conversation of statistical properties during the logarithmic data transformation.

Arcone, S., Spikes, V., Hamilton, G., & Mayewski, P. (2004). Stratigraphic continuity in 400MHz short-pulse radar profiles of firn in West Antarctica. Annals of Glaciology, 39, 195-200. doi:10.3189/172756404781813925

Arcone, S., Spikes, V., & Hamilton, G. (2005). Phase structure of radar stratigraphic horizons within Antarctic firn. Annals of Glaciology, 41, 10-16. doi:10.3189/172756405781813267

Arthern, R. J., Winebrenner, D. P., and Vaughan, D. G. (2006), Antarctic snow accumulation mapped using polarization of 4.3-cm wavelength microwave emission, *J. Geophys. Res.*, 111, D06107, doi:10.1029/2004JD005667.

Grima, C., D. D. Blankenship, D. A. Young, and D. M. Schroeder (2014), Surface slope control on firn density at Thwaites Glacier, West Antarctica: Results from airborne radar sounding, Geophys. Res. Lett., 41, 6787–6794, doi:10.1002/2014GL061635.

Konrad, H., Hogg, A., Mulvaney, R., Arthern, R., Tuckwell, R., Medley, B., & Shepherd, A. (2019). Observations of surface mass balance on Pine Island Glacier, West Antarctica, and the effect of strain history in fast-flowing sections. Journal of Glaciology, 65(252), 595-604. doi:10.1017/jog.2019.36

Morris, E. M. (2008), A theoretical analysis of the neutron scattering method of measuring snow and ice density, J. Geophys. Res., 113, F03019, doi:10.1029/2007JF000962.

Morris, E. M., Mulvaney, R., Arthern, R. J., Davies, D., Gurney, R. J., Lambert, P.,… Winstrup, M. (2017). Snow densification and recent accumulation along the iSTAR traverse, Pine Island Glacier, Antarctica. Journal of Geophysical Research: Earth Surface, 122, 2284 – 2301. https://doi.org/10.1002/2017JF004357

Nicolas, J., Vogelmann, A., Scott, R. et al. January 2016 extensive summer melt in West Antarctica favoured by strong El Niño. Nat Commun 8, 15799 (2017). https://doi.org/10.1038/ncomms15799

Scott, J., Smith, A., Bingham, R., & Vaughan, D. (2010). Crevasses triggered on Pine Island Glacier, West Antarctica, by drilling through an exceptional melt layer. Annals of Glaciology, 51(55), 65-70. doi:10.3189/172756410791392763

Yamamoto, J.K. On unbiased backtransform of lognormal kriging estimates. Comput Geosci 11, 219–234 (2007). https://doi.org/10.1007/s10596-007-9046-x