We thank both referees for giving up their time to review our paper and for providing constructive comments.

Our responses are below the referee comments in blue (bold and underlined).

Referee #2

1. I would like to see the argument for the timing mismatch between the model and the observations stronger and more coherently presented in the manuscript. Have the authors considered to use the high-altitude aircraft data during the outward and return flights to nail down this timing mismatch? It should be possible to see the turning of the winds occurring earlier in the model than in reality. Moreover, this allows a comparison between the nudged upper levels and observations, which is a more direct connection between the reanalysis forcing and the observational data. Also, is there additional data (from Rothera?) available that could serve to make the case of the authors stronger?

We have examined the aircraft data taken on the high altitude (around 3000m) approach and return legs. Unfortunately, given that there is only a time difference of 2.25 hours between the two legs and the fact that this gives are only two datapoints, it is quite difficult to say for sure whether there is a timing discrepancy between the model and reality. What is really needed is a longer term timeseries, which is difficult to accomplish at high altitude. Besides this, the differences involved are actually quite small – e.g. 9 hours is not a large amount of time relative to the duration of the event and the amount of turning of the upper level winds over the relevant period is fairly slight (around 37 degrees). Despite this, the data suggests that the modelled pressure was lower than that observed and the wind direction directed more towards the south, which is consistent with the timing mismatch suggested in the manuscript. Surface pressure data from Rothera shows a similar overall drop in pressure between the model and reality with similar timing (wind data is highly variable and not likely useful due to terrain effects and low level blocking). However, the nature of the changes are quite different with Rothera showing a constant pressure followed by a sudden drop and the model showing a gradual change, which may be indicative of errors in the large scale meteorological fields. Overall, though, it is hard to provide firm evidence that such a timing mismatch in upper level winds (and in the shifting of the low pressure systems) actually occurred. We have described this evidence in the updated manuscript and made it clear that the idea is likely to remain speculative, but plausible, and that there may be other causes for the mismatch in the timing of the low level winds as diagnosed by the AWS comparison. The revised text reads as follows (highlighted) :-

The moving eastwards of the small low pressure system over the ice shelf seen in Fig. 16a and b looks to be related to the movement of the larger low pressure system over the Ronne Ice Shelf (as seen in Fig. 2). It is possible that this system shifted prematurely in the model compared to reality and was responsible for the influx of southerly winds onto the ice shelf giving rise to the earlier change in 10 m wind speeds and direction compared to the AWS. Figure 2b shows that the movement of the low pressure system has resulted in the winds on the west of the Peninusla shifting so that they no longer impact perpendicularly to the ridge. It seems likely that this may have caused the cessation of the fõhn jets since fõhn flow generally requires winds that are close to perpendicular to the ridge. If the winds shifted early in the model compared to reality then this may have also caused the early cessation of the fõhn jets.

However, it is difficult to ascertain for sure whether there was a timing discrepancy between the model and reality for these large scale systems. Wind data at upper levels (above the mountain ridge height) would be useful for this since the flow is likely to be less variable and hence more representative of the larger scale situation. Unfortunately, only brief observations at such altitudes are available. For the aircraft observations made above the ice shelf at around 3000 m, the eastward flight leg (the earliest leg at around 20:07 UTC) and the westward leg (22:23 UTC) were only 2.25 hours apart, whereas what is ideally needed is a longer term timeseries. Comparisons with the model at the time of the earlier leg do show that the model pressure was 1.8 hPa lower than the observed mean over the leg and the wind direction was around 20° too low. These are both consistent with the upper situation changing too early in the model since the modelled pressure was dropping and winds rotating towards the south in the model. However, given the small margins involved it is likely that instrument uncertainties could also account for these differences. Comparisons to the surface pressure timeseries at Rothera (not shown) reveal a similar decrease in pressure between the model and observations after 0 UTC on 6th Jan, with no clear evidence of a timing issue. One difference, though, is that the observed pressure drops in a "step change" manner between 0 UTC on 6th Jan and 0 UTC on 7th Jan with fairly constant pressure in between, whereas the modelled change is more gradual.

Thus there is some evidence that there are discrepancies with the pressure systems and upper level winds of the model compared to reality. This would point towards a lack of accuracy with the large scale analysis that drives the model boundary conditions and upper level nudging, which in turn may affect the fõhn winds. However, given the evidence available, this is fairly speculative and it is possible that there were other causes for the timing differences seen in the low level winds between the model and observations. It should also be borne in mind that the change in upper level wind direction over the period during which the jets ceased was quite small; the wind direction was 237° at 6 UTC on 6th Jan and reduced by only 37° by 12 UTC on 7th Jan. Thus the margins of any error in the analysis are likely to be small, although the results here suggest that such small upper wind direction changes may be important for determining whether fõhn flow occurs or not. Also, a timing difference of approximately 9 hours is fairly small given the overall timeframe of the existence of the jets.

In summary, there are some differences between the model and the observations, but overall the agreement is good and gives confidence that the modelled jet behaviour was similar to reality in many aspects.

2. The quality of the surface energy budget analysis is somewhat hampered by deficiencies in the WRF surface scheme. The authors mention the unrealistic values for the longwave emissivity and the shortwave albedo of snow. In addition, the particular model treatment of the turbulent fluxes (especially with the lowest model layer at ↑27 m above the surface) may also explain why the modelled amplitude of the turbulent fluxes are smaller in magnitude than the fluxes presented in King et al. and Munneke et al.

What are the roughness lengths for momentum, heat, and moisture in the model? Also, it is conceivable that the energy balance fluxes (most notably the ground heat flux) is influenced by the initialization of the snow. Is the snow represented by a single layer? Or multiple layers? How is the snow initialized at the start of the run? Could the ground heat flux be influenced by the setup of the snow model part?

The selection of surface layer and land-use scheme is based upon the thorough testing and subsequent modification of the various available WRF schemes in order to determine those that best matched observations over ice covered surfaces, as detailed in Hines and Bromwich (2008).

The surface layer scheme used is that from the Eta model, which is based on Monin-Obukhov similarity theory, but with modifications following (Janjić, 2002). The roughness length for momentum is 10^{-3} m and the moisture and thermal roughness lengths are scaled from this following Zilitinkevitch (1995) as a function of the molecular viscosity for momentum and the friction velocity.

The snow pack is represented using four layers through the use of the Noah land surface model with modifications to deal with deep snow-packs described in Hines and Bromwich (2008). The density, heat capacity and heat conductivity of the snow-pack are based upon observations of Antarctic snow firn. So, the representation in the simulations presented in our paper were probably the best possible for ground heat flux calculations with the WRF setup as it was at the time. Of course there will certainly be scope for improvements, particularly regarding the tailoring of the scheme to the specific region of the simulation. Unfortunately, this is beyond the scope of our study.

The point about the initialization of the subsurface snow temperatures made by the Referee is well taken and is likely to be the largest area of weakness for the representation of ground heat flux calculations. The values provided within the WRF domain setup utility were used, which are based on annual averages. This therefore may introduce some errors in the ground heat flux and melting calculations since the use of seasonally varying subsurface temperatures tailored for Antarctic ice shelves would be more appropriate. Also, there may be some spin-up period for the temperatures of the sub-surface layers associated with the use of this data.

We have added to the discussion on these two issues in the revised text.

RE sensible and latent heat fluxes:-

Fõhn jets are also warm (near surface air temperature $> 0°C$) and so caused an increase in the amount of downward sensible heat flux at the surface. However, because the jet air is also dry, surface energy loss due to snow ablation (latent heat fluxes) tends to cancel out a lot of the surface heating effect due to sensible heating. This was the case in the modelling in this study and this is also consistent with the aircraft observations and AWS analysis mentioned above. However, the comparison to those results suggests that the sensible and latent heat fluxes were underestimated in the model, indicating deficiencies in the model representation of these processes and their link to the jets, or of the fõhn jets themselves. This is likely to implicate the surface layer scheme parameterization. The selection of the Janjić Eta scheme (see Section 2.2 for details) used in this study was based upon the thorough testing of the various available WRF schemes in order to determine those that best matched observations over ice covered surfaces, as detailed in Hines and Bromwich (2008). However, improved accuracy could likely be obtained through the use of roughness length values and scalings that are tailored to the Larsen C Ice Shelf.

RE ground heat fluxes:-

Whilst the model treatment of the thermal properties of the sub-surface snow pack were specially modified to deal with deep snowpacks, including the use of density, heat capacity and heat conductivity values taken from observations of Antarctic snow firn (Hines and Bromwich, 2008), it is likely that some deficiencies still remain. The values provided within the WRF domain setup utility were used for the initialization of the sub-surface snow temperatures, which are based on annual averages. This therefore may introduce some errors in the ground heat flux and melting calculations since the use of seasonally varying sub-surface temperatures tailored for the Larsen C Ice Shelf would be more appropriate. Also, there may be some spin-up period for the temperatures of the sub-surface layers associated with the use of this data. Therefore, it is recommended that sub-surface temperature data from longer term runs (i.e. with fully spun-up sub-surface temperatures) of this region are used for future studies (e.g. data from the Antarctic Mesoscale Prediction System, known as AMPS, or other polar WRF runs). The provision of sub-surface melt layers may also lead to better model accuracy in melting estimates.

==

Minor issues:

These have all been attended to, except where noted below.

p 5772 - I find the abstract rather long in its present form. Can the authors have a critical look at it and see which information may not be so crucial for the abstract after all?

We will take a look at this and shorten it.

p 5774 l.8: warming -> rising
p 5774 l.16: gives -> give
p 5780 l. 17: You state that temperatures higher than 0C in the cross-ridge flow would allow for surface melt. This is quite a general statement. It is also possible that there is no surface melt, for example if there is a strong inversion, or a high-albedo surface. Whether the surface is melting depends on the surface energy budget, not only on the temperature in the jet. Conversely, there could also be a melting surface if the air temperature at 250-350 m was below 0C. I suggest rephrasing to something like "The effect of these warm jets on surface melt is investigated in section 4."

This sentence has been changed to "Such temperatures could promote melting of the ice surface;" in order to indicate that the melting is not certain.

p 5780 l.22: 4 -> fourth
p 5784 l.5: should there be a reference to figure 9a here?
Rather, this should be Fig. 5a since we are referring to the time of 12UTC.

p 5784 l.12: figure 7b -> figure 9b (?)

Again, this should be Fig. 5b since we are referring to the time of 12UTC. Figure 9 is referred to in the next section.

p 5788 l.8: moving eastwards -> eastward movement
p 5788 l.16: movement eastwards -> eastward movement
p 5788 l.17: Peninulsa -> Peninsula
p 5791 l.10: Figure 15a and 15b shows -> Figures 15a and 15b show
p 5791 l.12: windspeed -> wind speed
p 5798 l.1: I find this a somewhat difficult statement. First, a shift of _9 hours makes that there is a shift of the turbulent fluxes with respect to the radiative fluxes (the latter are bound to the time of the day whereas the former are bound on the wind conditions). Second, whether the modelled effects of the jets on the ice-shelf surface are realistic entirely depends on the surface scheme in the model. Later, the authors acknowledge that this scheme is not fully suitable to study the surface energy budget.

We acknowledge that the diurnal timing issues are likely to cause different interactions in the model compared to reality and that a lot depends upon the realism of the surface scheme; the paragraph has been changed to:-

The good match between the model and observations presented so far give confidence that the development and evolution of the modeled jets are similar to that of the real jets, which might suggest that the modeled effects of the jets on the ice shelf surface will also be realistic. However, we also acknowledge that the interactions between the jet dynamics and the radiative fluxes will be somewhat different from those in reality due

to the timing issues described earlier. Also, the modelled impact of the jets upon the ice surface will be dependent upon the surface scheme of the model, which is discussed later.

p 5801 l.1: There are more possible causes than the reduced wind in WRF. It could be related to the surface scheme, and to the coarse representation of the boundary layer in WRF, with the lowest atmospheric level at _27 m above the surface. Can the authors expand on alternative explanations for the representation of the turbulent fluxes in WRF?

We have added a sentence here to mention that deficiencies in the model parameterization of the surface layer turbulent fluxes may also be to blame and referred to the discussion section for more details (as described above).

p 5807 l. 1: patter -> pattern

p 5812 l.5-7: This sentence is rather complicated, and not easily understood by nonnative speakers. Please simplify your message.

This has been changed to:-

interpreted in terms of Larsen Ice Shelf surface melting. Our results suggest that it is not likely to be the case that reduced upwind blocking, due to wind speed increases or stability decreases, will always lead to an increased likelihood of föhn events over the Antarctic Peninsula, as suggested in previous studies. Thus, increased westerly p 5823 fig.7: The labels A, B, C, D are not well visible. Please enhance the contrast between the blue background and the black labels.

The contrast has been improved.

p 5835 fig.19: I appreciate the attempt to plot all fluxes on the same vertical axis, but this looks a bit artificial to me. Would it be possible to define an anomaly from the latitudinal mean for each flux? It will lead to almost the same graph but the definition for each line would then be the same. All lines will be averaged around 0 by definition. Possibly, you could add the latitudinal means for the fluxes in the legend or as text in the figure.

We agree that this plot is unusual, but feel that what is suggested here would not be that much different from its present form. The mathematical definition would indeed be the same for each line. However, the means for each line would have to be listed in the legend, just as the values at the reference location are now. The disadvantage would be that it would also make the lines that are currently not adjusted harder to interpret.