Comments on reviewer #1, the original review is given in black, and author comments in **blue.**

Review of Hofsteenge et al. 2022 The surface energy balance during foehn events at Joyce Glacier, McMurdo Dry Valleys, Antarctica

The impact of foehn events on glaciers/melting has gained traction in the last decade, with many studies documenting the link between foehn events and melting in Antarctica, using a variety of methods. That being said, this paper contains novel aspects: this glacier is largely unstudied, especially in the SEB and foehn field, and reveals a more complicated story of the mechanisms behind foehn in this region, thereby progressing our knowledge. The study investigates a 14-month period, using both observations from an automatic weather station and output from a high resolution model. Perhaps some of the highest spatial resolution output used for foehn studies (in this region) thus far. The authors employ a combination of previously developed methods, highlighting an overall good agreement between methods, and between observations and model. A SEB model is used to look at the impact of foehn on various components of the energy balance as well as melting. The authors are aware of the limitations of their study and don't oversell their results or conclusions. They could actually highlight further some aspects of novelty which are only mentioned in the discussion. There are some areas of clarification required for the method, but I don't expect this to change any results or conclusions. The paper is structured well and is nice to read, with some useful and easy to interpret figures. Overall, I suggest **minor** corrections and outline these below.

We would like to thank you for taking the time to review this manuscript. We are happy to read your positive comments and to incorporate the suggestions you made to improve the paper. We describe below in more detail which changes have been made to the manuscript.

Page 2, Lines 29-38: Is it possible to include more details on the map of the Antarctic in Figure 1? You mention a number of mountain ranges and glaciers throughout the paper, but people unfamiliar with the area will find it difficult to put these into place. This is especially important when you discuss the interaction of the low pressure systems and airflow for the two different foehn mechanisms. I would suggest either 4 panels, with one providing more details (e.g. Transantarctic mountain range, Taylor glacier), or exchanging the AWS plot with an additional map.

We agree that Figure 1 could be more informative with the locations that are mentioned in the paper, therefore we made a new figure that includes all the locations and topographic features that are discussed in the foehn description.

Page 2, Line 34: Include a brief description of what a foehn wind is. TC is quite an interdisciplinary journal, but still read mostly by glaciologists who might be unfamiliar with the more atmospheric terms.

We added a sentence describing foehn winds: 'Foehn winds typically occur through topographic modification of flow in the lee of mountain barriers, resulting in strong and warm winds.'

Page 3, Line 57: Similar to above but for sensible heat, or perhaps just include *(through warmer and gustier than average winds)* after the words 'sensible heating'.

We have included here: 'through an increase in turbulent mixing resulting from gustier and warmer wind conditions.'

Introduction general: I think there are some aspects of novelty that you could highlight more in the introduction. Authors should double check that the following statements are true before including them, but if so, you could include them in your introduction. Is the 1.67km horizontal resolution of AMPS the highest spatial resolution that has been used for foehn studies in the Antarctic so far? Elvidge et al (2015) and Turton et al. (2017) have stressed the importance of higher resolution modelling for foehn winds due to their complex interaction with the topography. But most AMPS studies I can think of use 5km or coarser resolution for foehn studies. The higher resolution doesn't necessarily imply better results though, as the wind direction problem could be associated with this. You mention in the discussion that this is the first time a 4-component radiometer is used in the MDVs. Which was important for your surface temperature calculations, which are also not often included in foehn studies.

Thanks for your suggestions. Regarding the AMPS resolution, the mechanism resulting in foehn effects in the northern MDV has been studied with Polar WRF at a higher resolution than AMPS (0.5km nested grid scales). We have added in the introduction the following statement to clarify what model resolution has been used before: 'Steinhoff et al. (2013) used the polar version of the Weather Research and Forcasting (WRF) model on a 0.5km nested domain to study physical processes responsible for foehn events in Taylor, Wright and Victoria Valley.'

We did include as you suggested in the Introduction that a 4-component radiometer was used for the first time for a SEB study in the Dry Valleys, which improved the simulation of the surface temperatures. The revised text is: 'This is the first SEB study in the MDV to use a glacier based 4-component radiometer, which allowed for calibration of the simulated surface temperatures and increases accuracy of the simulated energy fluxes.'

Page 4, Line 90-95: When you list your observations and calculations, you don't mention surface temperature.

To clarify this we have moved the sentence that Tsfc is calculated from outgoing longwave radiation up to this section describing the observations.

Page 6, Line 137: Include a reference for the 12- hour spin up and removal. There can be quite a debate about the length of time that should be used for spin up, so a reference stating that this comes from the operational use of AMPS (and therefore isn't your decision) would strengthen the sentence.

Since the full AMPS output is not available anymore in an online archive, we have used a thinned AMPS product that was saved as 12-24h forecast hours. We have now cited a paper using AMPS in the same way (Seefeldt and Cassano, 2012), who provided us these AMPS products.

Page 6, Line 150: What resolution was the AMPS that Speirs used? On line 159 you explain the wind direction difficulties and therefore use a different direction threshold for AWS and AMPS. Was this also done by Speirs, or is this your decision? Does the choice of a different threshold affect when foehn is detected, and how did you settle on the specific directions you use?

In the study by Steinhoff et al. (2014) the same wind-direction criteria were used on AWS and AMPS output from a 2km nested domain. AMPS used in this study has a higher resolution and we think that the difference in wind direction between AMPS and AWS observations is caused by the more complex valley system around Joyce glacier compared to the wider valleys studied in Steinhoff et al. (2014). We have now clarified why a different wind direction threshold was used in the explanation of the Speirs method as follows: 'A different wind direction criterion is used on AWS data compared to AMPS forecast, since the weather station was located close to the valley side and AMPS does not capture well the topographic modification of the winds (Sect. 3.1)'. The wind-direction thresholds were chosen to represent the expected foehn wind direction and to make sure that it does not include the sea-breeze wind direction.

Page 7, Line 184: Include in this sentence that these criteria are applied to AWS near-surface data.

We have applied the near-surface conditions of Turton Part 1 as well using AMPS forecast for testing purpose. Under your comment below on 'Method general' we have included how we have made this clear in the manuscript.

Page 8, Line 190: Include that these criteria are applied to AMPS data.

Thank you, we have clarified this as follows: 'Sufficient isentropic drawdown, determined as follows from AMPS forecast (Fig. 2):'.

Method general: You need to provide more clarity on your foehn criteria that you employ after the evaluation stage and on what criteria you apply to near-surface conditions from AMPS and which to upper-air data (isentropic detection). For example, did you apply the methods to detect near-surface changes (Speirs, Wiesenekker and Turton Part 1) to AMPS nea2r-surface data and AWS data? If so, is this appropriate, given that the Turton Part 2 method should be used for AMPS data and not Part 1 method? AMPS was shown to relatively poorly represent near-surface conditions in foehn, and therefore a specific algorithm for the isentropic drawdown was used in AMPS. Or did you apply the Speirs and Wiesenekker method to AMPS near-surface and the Turton Part 2 method to AMPS upper air conditions? This becomes important in Page 10, line 231 where you compare the number of foehn in AWS and AMPS. Turton method shows the biggest difference in AMPS and AWS and you put that down to weaker humidity drops in AMPS, however, the original Turton method didn't look at the near-surface conditions in AMPS, but rather the isentropic drawdown to define a foehn. It's fine for you to adapt the methods for your own use/different location, but I would make it clear that what exactly you extracted from AMPS and how you applied the criteria.
Did you look at wind speed bias between AWS and AMPS to assess whether the 5 m/s Speirs and 4 m/s Wiesenekker is appropriate in AMPS if there was a considerable bias?
You say that at least 2 out of 3 methods must detect foehn for you to use it – but is this in AWS only data, or AWS and AMPS data? If just AWS, it needs to be a little clearer that the AMPS results are only used for evaluation.

Thank you for your extensive feedback here, that is very helpful. We have made several adjustments to clarify when AWS and AMPS data are used and for which purpose. We have used as in Turton 2018 Part I on near-surface conditions in combination by Part II on AMPS upper air forecast. We tested all methods using both near-surface conditions based on AWS observations and AMPS forecast. We have clarified that the Turton method was designed to combine AWS observations and AMPS forecast for the atmospheric structure in the description of the method: 'Turton et al. (2018) developed a more complex classification scheme where a combination of criteria to detect observed near-surface foehn conditions and AMPS forecasted foehn characteristics in the upper-air are used for foehn detection over the Antarctic Peninsula.'

We clarified that we test the method here as well with AMPS near-surface conditions in Section 2.4.3 as 'For evaluation purposes, this method is also tested using AMPS forecasted near-surface conditions in Sect. 3.1' . In Section 3.1 we have clarified again that AMPS near-surface conditions are only used for testing and in the remainder of the manuscript AWS near-surface conditions are used with the following two sentences: 'We applied the 3 methods described in Sect. 2.4 to identify foehn conditions at the Joyce Glacier using both near-surface conditions based on AWS observations and AMPS output for the 14-month study period.' and 'Throughout the rest of this paper a combination of the detection methods using near-surface conditions from AWS observations are used to classify foehn.'

Regarding AMPS bias of wind direction, we found on average a slight negative bias in AMPS, which is due to a negative bias under calm conditions (poor representation of down-glacier winds, Sect 3.1). During strong wind events, AMPS often has a positive bias and therefore we think the AMPS windspeed bias does not impact the foehn detection.

Foehn detection and model evaluation
Some comments from this section are include in the above paragraph.

Page 11, section 3.2: Is there a reason you decided not to use the AMPS SEB output for an additional analysis, such as in King et al. (2015, 2018, and others)? I was occasionally confused by the word 'simulated' for the SEB section, as it comes after the use of AMPS for many paragraphs. The SEB model is run using observations, so it is more observationally forced than simulated. I wonder if the word 'simulated' should be linked only to AMPS to reduce this confusion. I'll leave this up to the authors however, as perhaps my previous use of SEB output from AMPS has skewed the way I am reading this.

To solve the confusion around AMPS and SEB simulations, we followed the suggestion by reviewer #2 to change all references to the AMPS data by using 'forecast' and using 'simulations' for the SEB results. We have considered to use the AMPS SEB output to discuss spatial SEB patterns, however we found that it poorly represented the SEB of the dry valley glaciers. The limitations of AMPS are mainly that i) the snow model does not allow solar penetration, ii) AMPS solar radiation poorly captures the variability through cloud cover and iii) the sensible heat flux did not capture the typical unstable conditions.

How did you create the ensemble? How many runs is this and what initial variable or value did you alter?

The ensemble was created using the combinations of parameter values listed in Table 1. To clarify we have added a reference to this table when discussing Figure 4 showing the ensemble spread.

Page 12, Line 285: Is the anomaly plot from AMPS? Include in figure caption or text to make it clear.

This has now been included in the captions of figures showing AMPS forecast output.

Page 13, Line 291: Some important places for context here but no map to point them out.

We have added a reference to Figure 1, which now is adjusted to show all the important places (see answer to previous comment).

Page 14, Line 305: If you applied the Turton Part 2 algorithm to AMPS data, to detect isentropic drawdown, this could be the reason for a higher number of foehn events using this method than the other two. If Joyce Glacier is more susceptible to isentropic drawdown, the Turton method is more likely to pick that up than the methods using only near-surface characteristics.

Thank you for this suggestion, this could indeed be the case. It might be that Joyce is occasionally influenced by drying & isentropic warming, while windspeeds do not remain elevated and are therefore not picked up by the other methods.

Page 14, line 315: What is the *(increasing)* for here?

Changed into 'March and October can be considered as transition months with a decreasing or increasing contribution of solar radiation', to clarify in March solar radiation decreases and in October it increases.

Page 14, line 325: This sentence is a little confusing, you write both simulated and observed surface temperatures twice, perhaps shorten the sentence or remove the 'based on both simulated and observed surface temperatures'?

We have rephrased this sentence to simplify and clarify it, as follows: 'Although simulated surface temperatures have a positive bias in summer (Sect. 3.2), both simulated and observed surface temperatures are often warmer than the summer air temperature, confirming the often negative sign of SH'.

Page 15, line 339: Do you detect/categorise katabatic winds too, or you just look at the main characteristics to decide that during non-foehn conditions, the SH flux is due to katabatics? Can we be sure that we are not falsely categorising katabatics as foehn winds? Or vice versa? Especially with the Wiesenekker method, which doesn't take into account temperature or relative humidity.

We believe that we do not falsely detect katabatic or other non-foehn wind events, since we use a combined method where two out of the three methods need to flag the 6h period as foehn. In practice this means that either the Turton method or the Speirs method confirms the isentropic drawdown or the surface warming & drying characteristic of the foehn event. We realize that adiabatic warming of katabatic winds could result in a similar temperature and relative humidity signature. However, the wind events are unlikely to be of a katabatic nature as they are associated with cyclone activity and synoptic flow that leads to foehn events rather than air originating from the Antarctic plateau. Please note that we have changed the structure of the Discussion & Conclusion section (as suggested by reviewer 2) and have included more discussion around the katabatic vs foehn origin of these winds here to clarify this point.

Page 16, line 352: Is this statement about no MDV melt during cloud cover your result (so specific to Joyce Glacier during the 14 month period) or is this information from elsewhere and more of a generalisation. As figure C implies your data alone, but by saying 'MDV glaciers' it seems like you have more evidence for this statement.

We rephrased this statement, to clarify that previous research has shown that MDV glacial melt is solar radiation driven and that in this study we find that melt does very rarely occur during cloudy conditions, as follows: "Melt at MDV glaciers is driven by solar radiation (Hoffman et al., 2008) and melt does rarely occur during days with clouds in the 14-month record at Joyce (Appendix C, Fig. C1). 83% of the surface melt at Joyce Glacier occurred during clear-sky conditions (Ne < 0.2), and only 0.5% during overcast conditions Ne > 0.8).".

Nice figures and use of space on the figures.

Page 17, line 365: There has been a study that looks at the longer-term effect of foehn warming on the snowpack, or pre-conditioning of the surface for future melt. E.g Kuipers Munneke et al. 2014, Elvidge et al. 2020. Could include reference to similar studies also on page 23, around line 505.

Thanks for these suggestions. We have included the following references on the pre-conditioning effects found by these studies in the Introduction and Discussion: 'Here, foehn events cause the highest melt rates and the increased melt associated with foehn plays a preconditioning role in ice shelf melt and instability (Kuipers Munneke et al., 2014)' and 'However, in a similar way foehn plays a major role in prolonging the melt season at Larsen C ice shelf and spring foehn events precondition the ice shelf for enhanced summertime melt via reduced albedo (Elvidge et al., 2020)'

Page 17, line 377: include *'during such conditions'* after '*...key role in melt occurrence*'.

Thanks, done.

Page 18, line 379: Did you look into the melt immediately after foehn events? 6-12 hours after the foehn event for example? It could provide a nice extra result.

We have had a look at the diurnal cycle in energy fluxes and melt rates on the days before and after days with foehn, as from Figure 9 is visible that some of the days with melt occur after foehn events. Looking at a few hours after a foehn event will be biased by the diurnal cycle as melt solely occurs when solar radiation is peaking. We did not include these diurnal cycle plots on the days after foehn since it was hard to draw conclusions from them based on this short study period where also days after foehn overlap with days before the next foehn event. This idea that large melt rates might occur after foehn events will be considered in follow-up study using a longer term foehn-SEB record.

Page 18, line 382: In the brackets for the figure citation, include that the reader should look for the grey line on 9a, as it took me a while to see it amongst the other colours and bars.

Thanks, we have added this.

Page 19, line 397: include '*in which*' after '*68% of the hours*'

Thanks, done.

Page 19, line 398: change *'is peaking'* to '*peaks*'

Thanks, done.

Page 19, figure 10: Consider changing the colourbar to one that isn't rainbow, as this can be difficult for those with colourblindness. Interesting way to represent wind direction – I like it! Same for figure C1.

It was difficult to find a colourblind-proof cyclic colormap, therefore we changed it into wind vectors as suggested by reviewer 2. We adjusted the colour palette of the line graphs in the manuscript as well to make them more colourblind-friendly.

Page 20, line 410: is this steady south-easterly a product of your foehn algorithm though?

This can indeed be partly caused by the foehn algorithm. However, we found that running the Speirs algorithm without the wind-direction criterion gives very similar results, suggesting that this constant wind-direction is typical for foehn conditions and a clear deviation in summer from normal conditions where a diurnal wind regime occurs.

Page 20: first paragraph of section 5: Some of this novelty should be clearer in the introduction.

We have added the following sentence in line with your previous suggestion around the radiation observations in the Introduction: 'This is the first SEB study in the MDV to use a glacier based 4-component radiometer, which allowed for calibration of the simulated surface temperatures and increases accuracy of the simulated energy fluxes.'

Page 21, line 453: You also had quite a short duration of data, so could not look at the interannual variability or be sure that this 15% value is characteristic for your region either.

We agree with this, we can only say it is within the inter-annual variability of the northern MDV foehn. We clarified this in the manuscript by adding: 'The average foehn occurrence found for the 14 month study period at Joyce Glacier might not reflect the climatological mean, but it lies within the large inter-annual variability of the northern MDV (Speirs et al. 2013)'.

Page 21, line 460: The relatively coarse vertical resolution of AMPS (44 levels) may also not allow simulation of conditions closer to the surface. Many regional models now have 60-70 vertical levels.

That's right. Some recent studies use even higher resolution models to study foehn events (e.g. Umek et al 2022) and show that vertical resolution improves the simulation of cold-pool and foehn interaction and therefore the foehn signature at the surface. We realize that AMPS might not fully represent the conditions close to the surface and therefore we use AWS observations for the near-surface conditions and use AMPS solely to study the larger-scale atmospheric structure during foehn. We have added a statement around the vertical resolution in Section 3.1: 'Secondly, an accurate representation of near-surface conditions might be limited by the vertical resolution of the AMPS forecast'.