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.

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.

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.

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'.

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.

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

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.

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?

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

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

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 near-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.

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.

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

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

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

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.

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

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'?

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.

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.

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.

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

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.

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.

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

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

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.

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

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

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

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.