Dear Editor,

We thank the reviewers for their helpful and constructive comments and respond to them below. The manuscript was revised accordingly. There is new content in the abstract and in the introduction and the results section contains a new figure. On the other hand, some redundant content was removed. We improved the structure of the manuscript by adding subsection headings and splitting long paragraphs into shorter ones. The first subsection in the Results was moved to a separate chapter describing the acquired data, which is now also subdivided into subsections. Each subsection in the results now starts with one sentence summarizing what was done and why this analysis was carried out. We hope that these measures help to make the manuscript less confusing for the reader.

Referee comments are in black, answers in blue.

Report 1, Simon Filhol

Investigation of a wind-packing event in Queen Maud Land, Antarctica by Sommer et al. presents a set of unique and novel data acquired with state of the art instrumentation. This is the first time, to my knowledge, that the combination of such detailed spatial coverage of the snow cover morphology and hardness evolution of the snow surface have been acquired simultaneously. Snow hardness is thought to be an important component of the erodibility of snow, therefore having a control on where snow is being removed at a small scale (decimeter to decameter), and having a control on fluxes of snow moved by wind at a larger scale (hectometer and larger) (Li and Pomeroy 1997). The data themselves should be of high interest to the scientific community.

Over the span of almost a month, the authors collected 9 laser scanner scans, and 454 snowmicropen profiles at 12 different dates. During this event, it is clear that there is one 10cm snowfall followed 9 days later by a drifting event (as captured by the SPC sensor). The lidar as well as the SMP data brackets both events quite well. While there have been laser scanner records of the snow surface transformation during storms in Antarctica on sea ice and on the ice cap (Picard et al. 2016, Trujillo et al. 2016), this study adds a new aspect to the system with the snow hardness measurements.

I recognize the effort of the authors to revise the manuscript following a first round of reviewers requesting major changes, but I find the manuscript still in need of in-depth changes to become clearer to the reader. In the response to the reviewer, the authors mention the original intent of this manuscript to be a letter, which, in my opinion would require even further clarification of its structure. Many parts of the text did not or only slightly changed. Section 2 and 3 are poorly defined when reading the text, and present data not used in the analysis. Many syntax errors add to the confusion of the reader. As a result, the assertions and deductions done in the discussion are poorly convincing. On the opposite, the addition of the lidar error estimate is very useful, as well as the set of new figures is more relevant, compact, and complementary to the text than the previous version.

Thank you for your thorough review and many helpful comments which helped to improve the manuscript. Section 2 is now divided into subsections and some of the longer paragraphs were subdivided as well. Some of the content was reorganized to make things clearer, e.g. the figure showing an overview of the study site was moved to the beginning to this section.

The long first subsection in the Results (Overview of the investigated period) was moved to a separate data chapter and itself subdivided into subsections.

It's true that the new "data" section shows data which are not used in all parts of the study. We do not believe this to be a problem. On the contrary, we think it is bad practice to simply neglect unwanted data and to not clearly declare that and explain why this data was discarded. Furthermore, none of the data is completely discarded (see also comments below). Note that the measurement setup may feel unstructured at times, but the reason was that we had to flexibly adjust our measurement plans given the developing weather conditions. The syntax errors were corrected, based on the specific comments given below. Thank you for pointing them out.

The introduction could use some re-structuring with 1) a more extensive background for the reader, providing references to previous work on the topic, 2) a throughout paragraph on the impact of this study, and 3) a clearly articulated research question. After reading the introduction, it is still unclear where the manuscript is heading to, and what is the main question/hypothesis being asked/tested. The title of the manuscript eludes to a descriptive paper while in fact the reader finds a test of a statistical model (proven to not be useful at the end). So, are the authors presenting a statistical model to estimate the spatial distribution of snow hardness? Are they presenting an event based description of a phenomenon? Or, are they testing out a hypothesis?

The introduction was reorganized along these lines and several new references and more details were added. The goal of the paper is to investigate the observed event with regard to all aspects related to wind-packing and to compare it to the wind tunnel experiments where possible. The paper is therefore descriptive as suggested by the title. The "test of a statistical model" (This probably refers to the correlation between the hardness and Sx?) is part of the comparison to the wind tunnel results.

Data and methods for the three types of instrumentation (lidar, SMP, and met-data) are presented all at once rather than independently. For instance, the last paragraph of this section starts by presenting the SMP measurements and finishes by explaining lidar error estimates. Moreover, much of the content here could be part of the results rather than methods.

Based on the reviewers comments, we restructured Section 2 better by introducing subsections. Note, however, that since the measurements are also partly interconnected (e.g. we're interested in the snow depth changes based on the TLS data at the SMP positions) it is difficult to present them independently. Before the SMPs are presented, we introduce the TLS measurements and their accuracy in general (inclination changes). After the SMP measurements are introduced, the TLS accuracy with regard to these measurements is presented.

The result section starts with a whole paragraph pointing to figure 4, with almost no other content. The actual relevant content related to this figure (where it should be referenced) is somehow split between the previous section and section 3.1. Further in this section, collection and processing methods are mixed with results (e.g. paragraph from line 24-35 on page 9). This makes it harder to combine all this information together. Also, many data are presented throughout this section and dropped out of the analysis. Why including and presenting them then? If the authors think that they are of any use for other purposes, then they could be organized and included in a supplementary to the manuscript.

As mentioned above, Fig. 4 and the corresponding description of the data was moved to a separate data chapter and better structured.

As to presenting the unused data, we think that it is important to clearly state and explain why this was done. It's true that the explanations are quite detailed. This is because it was not straightforward to decide which SMP measurements can be kept and which cannot be used in detail. Furthermore, none of the data is completely unused. All SMP measurements are shown in Fig. 4D. This is important because it shows, for example, that the range of the SMP hardness is very high on most days. If only the SMPs shown in Fig. 6 were presented, this result would be completely lost.

Section 3.4 and 3.5: The authors, opportunistically present data related to a barchan dune starting on page 13 with little mention of this in the introduction, and how it relates to the rest of the analysis. These data are unique though! The barchan dune is loosely defined, and it appears to be a made of at least two barchans. Notice in figure 10 how the lower horn (on thelower left of the image) has itself two smaller horns. Bedforms are known to merge and split. This bedform could be a merge between at least two. Instead of looking at the difference between two scans, could the scan of January 11 itself show the bedform in a clearer manner than the DSM difference? Barchans in Antarctica are known to be more elongated (Kuznetsov, 1960, and Kotlyakov, 1966).

The surveyed barchan dune is now placed in better context in the introduction and it is explained why these measurements were done. Since all of the new snow was reorganized into barchan dunes during the main drifting snow event we had to look at one of them to study the hardness of the newly deposited snow. The formation of depositional features is wind-packing. Thank you for pointing out that this dune could be a merge of two dunes. We added this information in the manuscript.

The scans themselves show a less clearer picture of the dunes than the DSM differences.

When it comes to the hardness measurement on the barchans: - The scatter might be reduced by plotting the SMP hardness as a function of the radial distance from the crest.

We do not understand how this would work. Is the crest not a line?. How would the radial distance from a line be defined? In addition, the location of the crest is not clearly defined

- Is the trend more influenced by the date of the measurement or the actual position on the barchan? A GLM taking into account the date (i.e. hardness distance to tail + date) could help to detect if the date of the measurement plays a significant role or not into the correlation. If it does, why would this be the case? What processes could have come into play? Moreover, a Pearson's coefficient of 0.4 leads to a R 2 of 0.16 which shows almost no dependence of SMP hardness to the distance to tail.

We tested adding the "number of days since the drifting snow event" (Ndays) to the model. (3 for SMPs acquired on 3 January, 6 for those from 6 January, etc.) There is no significant relationship between Ndays and the hardness (Pearson's r of -0.25 and p-value of 0.06). A GLM with both parameters has an adjusted R^2 of only 0.14 and Ndays is not a significant parameter (Anova p-value of 0.4). We think that Fig. 6 establishes quite well that there is no clear relationship between the hardness and time in this dataset. A reason why Ndays is not significant in the extended GLM is also that distance to tail and Ndays are not independent parameters.

As to the low value of the correlation between the hardness and the distance to the tail: This is true but we clearly state in the manuscript that this result should not be overestimated.

- Is it possible to see in the SMP data the difference between the snow of the barchan and the underlying layer? If yes, showing the raw data as a section through the barchan would be very insightful to the reader.

The interface between the newly deposited snow and the old snow is unfortunately not visible in the SMP profiles. If that had been the case, we would also have used this information to help choose which SMPs to analyse in detail because this was basically about deciding which profiles have newly deposited snow at the surface. We added a new figure (Fig. 12) showing SMP profiles in one transect across the dune. This is now also discussed in a paragraph by this figure.

The discussion is also mixed. For instance, paragraph 2 compares the data to the wind tunnel experiment, paragraph 3 presents interpretation of the hardness data on the bedform, and paragraph 4 is again talking about the wind tunnel experiment. The discussion also contains contradicting assertions about the potential cause for hardening. The authors justify the trend in hardness of the barchan with the tunnel experiment, when afterwards, the model derived from the tunnel experiment is shown to statistically not hold for this dataset. Overall, this manuscript contains an interesting and unique dataset, but it would require some in-depth changes to be convincing, and clear to the reader.

Paragraph 2 discusses the necessity of drifting snow that was observed both in the wind tunnel and in Antarctica, and paragraph 4 discusses the differences of the Sx analyses between Antarctica and the wind tunnel. Because this is related to the barchan dunes, it is done after discussing the hardness variability on the barchan dune in paragraph 3. In fact, paragraph 4 continues to discuss the hardness variability on the dune, namely with respect to Sx, while paragraph 3 was about the distance to the tail. We do not see why this is a problem?

We also do not understand what is meant with "contradicting assertions". Where do we justify

the trend in hardness of the barchan with the wind tunnel experiment? We think there is no justification but only comparison between the two cases.

"Model derived from the tunnel experiment..." This probably relates to the correlation between the hardness and Sx that was observed in the wind tunnel? It's true that this relationship was not observed in Antarctica, and we explain that this is due to a lack of simultaneous measurements.

Specific comments:

Page 1, line 13: "Wind-packing and its results" the use of its results seems vague. What is specifically meant?

This sentence was rewritten and split in two. "Wind-packed snow has been described qualitatively in many studies especially in Antarctic literature
citep[e.g.][]Benson1967, Endo1973, Kotlyakov1966, Schytt1958, Seligman1936. These studies also suggest different physical processes but it remains unclear which of these processes actually happen during a wind-packing event."

Page 2, line 6: "The Antarctic event ..." Odd formulation, as if this was a widely recognized event, but also multiple events are mentioned before.

We changed this to "The wind-packing event observed in Antarctica..."

Page 3, line 10: What is the actual wavelength of the laser?

1064 nm. This was added to the manuscript.

Page 3, Line 14: 'the coreless winters' not sure what is meant by coreless.

This means that the temperature is quite constant during the winter months. The sentence was changed to "The winters are rather mild and coreless, meaning the temperature is almost constant during several months."

Page 6, line 4: "All (accurately known) SMP positions have a range below about 100 m" what is meant here?

This was indeed a bit unclear. We changed the sentence to "All (accurately known) SMP positions were located within a distance of 100 m from the scan position."

Page 7, line 11: 'very' not necessary.

Ok, we removed it.

Page 9, line 1: 'the logbook notes about 10 cm ...' is there a verb missing, or a miss use of the verb to note.

The sentence was changed to "It was noted in the logbook that there was about..."

Page 9, line 9: "each" and "furthermore" should be removed, or the sentence syntax needs to be reviewed

"each" refers to the fact that five measurements were done both at the top and the bottom of the new snow and "furthermore" refers to the fact that these measurements were done in addition to the measurements of the average density. We do not see the problem with this sentence.

Page 9, line 11: remove "very"

Ok, we removed it.

Page 10, line 1, 4, : remove all unnecessary "very"

Ok, we removed them.

Page 10, line 7: reason should read reasons; "in the following" could be replaced by "any further"

Both changes were incorporated.

Page 10, line 8: "At first glance" used with "appears to be" could be simplified. Remove "very"
The "very" was removed. The rest of the sentence was not changed for now. It is unclear to us how it could be simplified?

Page 10, line 9: the sentence has two verbs "was" and "reaches", and two subjects. Syntax problem.
We rephrased this sentence, but think that since there are two clauses, that two verbs and subjects are allowed?

Page 11, line 3: 'the positions ... are'
This was corrected.

Page 11: Many use of "therefore" where not actually needed.
One instance of "therefore" was removed another was replaced with "As a consequence"

Page 13, line 11: could read 'Fig. 9 shows the dune for each of the four days it was scanned ...'. The 'four scan days' sounds odd in this case.
We think the expression "scan day" is an easy way to say "day the dune was scanned". Writing this out would make the sentences more complicated to read.

Page 14, Fig9 caption. The expression 'scan day' is confusing and not quite accurate.
See comment above. It is unclear to us why this would be confusing and not accurate?

Report 2, Charles Amory

1. P7, L7 and elsewhere: in many instances "begin" is employed as a noun. I'm not a native English speaker but I think "begin" is a verb and "beginning" is the noun you actually want to use. If you agree please correct accordingly.
You're right of course. Thanks for pointing that out.

2. P9, L29 and elsewhere: Prefer 10 -3 N to mN.
We changed the unit from mN to N, but preferred 0.0xx N to the exponent notation.

3. P9, L31: remove "was"
Done. We also put "acquired on 18 December" in parentheses to make the sentence easier to read.

Regarding the length of the revised version, I think the abstract is a bit short and could contain more of the main and interesting results of the paper. I tried to list them below
- The SMP hardness increases after the main drifting snow event are significantly higher than anything achieved in the wind tunnel. most likely due to higher wind speeds and more intense drifting snow causing more compaction and hardening in the natural environment.
- Time and sintering are not the dominating processes in wind-packing, in agreement with previous wind tunnel experiments but in disagreement with previous literature, but the measured hardness variability could not be adequately explained with the available data.
- The field data exhibits a low correlation between the wind exposure Sx and the SMP hardness change, but simultaneous measurements of the hardness and Sx are needed for a direct comparison with the wind tunnel experiments

- The wind exposure, wind speed and drifting intensity at the moment of deposition are probably more important than the age of the deposition to explain the measured variability of the hardness

The Abstract was extended along these lines. Thank you for your comments and the endorsement for publication.

Report 3, Kouichi Nishimura

In general, it looks like my specific but minor comments are carefully taken into account. However, I am still concerned about the issues listed below and these should be satisfactorily addressed before the paper can be accepted for the publication.

First of all, contents are largely expanded particular in Sections 2 and 3 according to the referee's suggestions including myself. Contents newly added on the manuscript are probably based on the logbook recorded in Antarctica. In actual they are quite useful to recognize the situation during the observation periods. However, the descriptions are now too long and rather redundant. Since this is not a data-report but a scientific paper, authors need to set the focus on the specific issues. Redundant part should be eliminated and contents should be made much more straightforward. In fact, conclusions are deduced from the data obtained on group F only and others are not included in the analysis. Thus, detailed introductions about group A to E are not always necessary.

Thank you for reviewing this manuscript a second time and for your comments. In this case, we don't agree that the introductions of group A to E are not always necessary. The SMP data from the other groups is used as well. Group F contains only SMPs taken after the main event. The comparison "no drifting" to "with drifting" (Fig. 7) is not only based on group F and the other groups can therefore not simply be neglected. Other groups are also used in Fig. 6 and all data is shown in Fig. 4D. This is important because it shows, for example, that the range of the SMP hardness is very high on most days. If only the SMPs shown in Fig. 6 were presented, this result would be completely lost.

Indeed, not all data could be used in every part of the study. We think that it is important to clearly state and explain why some of the SMPs could not be used everywhere. We agree that the explanations are quite detailed. This is because it was not straightforward to decide which SMP measurements can be kept and which cannot be used in detail.

Further, if the authors would like to emphasize the points that drifting snow is necessary for wind-packing (but is not always sufficient) and subsequent drifting snow event increases in surface hardness, Figs. 6 and 7 will be fully enough to come to the conclusions. I do not think the following analysis and the devious explanations are needed.

We agree that in principle Fig. 7 is enough to show that drifting snow is necessary for wind-packing, but this is not the only thing we want to show in this paper. The goal is to investigate all aspects related to wind-packing of the observed event and compare this to the wind tunnel experiments. The formation of barchan dunes is the result of wind-packing. It just so happened that the main drifting snow event reorganized the new snow layer into barchan dunes. To measure the hardness of the snow deposited during the main drifting snow event we had to consider the structure of barchan dunes. The snow surface everywhere else consisted of old snow.

Although I can appreciate the efforts very much, unfortunately they did not work as were expected and all outcomes seem weak to persuade the readers. I can recognize quite well that the observations in the field, in particular under the harsh conditions in Antarctica, measurement conditions and possible observation period were extremely restricted. Such excuses are also found in the manuscript. Probably due to these limitations, the discussion parts are not straightforward and no distinct evidences are found out. We think that the result concerning the necessity of drifting snow is clear and that there is enough evidence to make this conclusion. On the other hand we find no correlation between the hardness and Sx. We explain in the discussion section why this may be

the case, and the conclusion that simultaneous measurements are necessary is also an important result.

It will be surely useful when the manuscript discusses the snow dune formations; the distributions of snow hardness, topography from tail to crest are finely observed probably for the first time. However, when the authors would like to investigate the wind-packing as shown in the title, the contents after the chapter 3.4 should be shortened largely.

As mentioned in the comment above, the formation of the barchan dune is the result of wind-packing. The dune was surveyed as a feature of newly deposited snow and not specifically as a barchan dune. For our analysis it would not have mattered if the drifting snow event had formed whaleback or other types of dunes. It's true that Fig. 9 showing the evolution of the dune after the drifting snow event is not directly related to wind-packing. However, we included this Figure based on comments in the first round of reviews and we think it's important to know that the dune did not change much in the timespan that the SMP measurements were taken on the dune.

Specific comments are listed below.

Figure 4: Date shown on the bottom of the figure should be clarified as we can recognize the date specifically, as is shown in Figure 6.

The length of the ticks corresponding to the labels was increased to make this clearer.

Page 9, line22: "Even after the main drifting snow event, there are many SMPs with very soft snow at the surface. This shows that drifting in itself is not a sufficient condition to form a wind crust". Would you like to say that the drifting snow is necessary condition but is not always sufficient condition? If it is the case, I strongly recommend to declaring like this, such as in the abstract and conclusion.

The fact that there was soft snow after the main drifting snow event does not really show that drifting snow is necessary but only that it is not sufficient. We therefore left this sentence as it is. That drifting snow is necessary is shown in section 3.3. We added a sentence there, clearly declaring this result.

Page 11, line 4: "therefore not further analyzed" Similar notes can be found repeatedly in the manuscript. When the accuracy of the data is not enough and you do not use the data in the analysis, I am not certain it should be fully declared. As is mentioned before, setting the focus on the measurements at group F looks better strategy.

As mentioned in a comment above, not only SMPs from group F are used and simply neglecting data for some parts of the study without explaining why is bad practice in our opinion.

Page 18, line 5: "The observed period" What do you mean by that? Duration of the observation period in the wind tunnel and in the Antarctica is the same? Perhaps it is not true.

This sentence refers to the sequence of events (snowfall, no-drifting, drifting) which was similar in both cases. We did not mean to suggest that the duration of the observed periods was the same. What is meant by that statement is also explained on page 18, lines 5-8, and also in the introduction on page 2, lines 4-6. (locations in the old version of manuscript) We also changed "observed period" to "observation period" to hopefully clear this up.

Page 19, line 8: "This is most likely due to higher wind speeds and more intense drifting in Antarctica. This leads to more frequent and more powerful impacts of snow particles on the surface causing more compaction and hardening" This is probably true, but there seems a gap in the argument. Preferably this speculation needs sort of proof quantitatively or even qualitatively.

We added a sentence about the increased momentum of the saltation particles. This hopefully clears things up.

Page 19, line 26: 'Time and sintering are not the dominating processes in wind-packing and is due to the impact of snow particles". I am just curious whether the authors have tried X-rays

CT analysis for the surface snow in the wind tunnel experiment. I believe it will give us the useful information to investigate the dominant process.

We did analyse some samples from the wind tunnel in the CT. It was observed that the density of the wind-packed snow increased and the SSA decreased compared to the initial new snow. These measurements were interesting but based on their (although rather quick) analysis we think it is difficult to gain knowledge about the physical processes from this data. We think that the combination of SMP and Kinect data is more promising in this respect.

Line 15: "We are not suggesting that wind exposure is not an important factor for wind-packing in Antarctica". Do you mean the effect of wind cannot be excluded? It actually appeared abruptly and I could not follow what leads this declaration. Is this inconsistent with the wind tunnel experiments in which the wind-packing was not formed without drifting?

Here, we refer to the effect of wind exposure and not the effect of wind itself (which clearly has an effect). The paragraph starting on page 19, line 30 is about the wind exposure and why the result was different in Antarctica than in the wind tunnel. So, we do not quite understand why this declaration was abrupt.