Response to Comments from Reviewer 1

AUTHORS: *We thank the Reviewer for their comments; edits based on their input have improved this manuscript.*

The authors present new elevation data collected using kinematic GPS in Antarctica and discuss their relevance toward validation of satellite laser altimetry data from the ICESat-2 mission. The paper was concise and well-written, and I am happy with publication almost as-is.
I have three major comments:

(1) While the authors discussed elevation changes associated with sastrugi migration, there was no discussion of other surface processes, primarily firn compaction in the context of

(a) Temporary (perhaps elastic) compaction of snow/firn from the weight of the PistenBully, which might not be captured from the track depth measurements. How heavy were the PistenBullys and is this effect negligible?
AUTHORS: The reviewer is correct: Our track-depth measurements almost certainly would not capture the elastic effect you are describing. PistenBullys weigh about 10,000 kg, spread over a 10 m length and 2 m width scale. While the elastic effect of that weight pushing temporarily on the firn might not be negligible, our suspicion is that the impact is below our leading error term, which is the track depth uncertainty and non-uniformity (~6 cm ± 1.5 cm). We also note that if the elastic effect is less than the precision of our GPS data (~4 cm) then we may not be able to observe this without creative survey methods. If extant, such an elastic effect would present as a negative bias of the vehicles relative to a non-invasive data set, such as airborne or spaceborne lidar. We intend to experiment with some creative survey methods, as time allows, at the start of this field season. But currently, this is beyond the scope of this paper. Further, we have other strategic plans to reduce the track-depth uncertainty; this includes running the GPS from the sled (which seemed to float at a consistent depth in the snow), as opposed to the PistenBullys (where the tracks seemed to dig into the snow at varying depths).

(b) Climate-driven firn compaction over < seasonal to multi-year timescales showing up as elevation differences between GPS- and Operation IceBridge-derived estimates. I think it could be useful if the authors included a time series of modelled elevation change from firn processes (data available at http://www.staff.science.uu.nl/~ligte104/DATA/) at one or more locations along this transect.
AUTHORS: The reviewer is correct that time variable climate-driven firn compaction could manifest as an elevation difference between our measurements and the airborne data presented here. The seasonal effect is less of a concern as the elevation response to climate-driven firn processes are effectively annual, and in this region should be driven by the annual cycle of temperature. As the UAF data were nearly coincident in time with the GPS data collection, the impact of seasonal firn processes on elevation should be minimal. The ATM data was collected with a larger offset with respect to time of year (~2 months) which may be impacting our comparisons. The secular trend in elevation owning to climate-driven firn processes is a thornier problem, as in this sector the climate drivers (accumulation rate, temperature) are poorly known. Both the secular and annual trends in elevation owing to firn processes will be a primary concern in our second paper on the topic, which will compare two seasons of ground-based data collection with airborne and spaceborne lidar data. We have added the following text to the document:
"We note that our analysis does not attempt to account for elevation changes due to the temperature- and accumulation-rate-driven effects of firn compaction (Li and Zwally, 2015). In this region, we expect variation in firn compaction rate to be driven by changes in firn temperature, which have a large seasonal amplitude and a much smaller secular trend. As the firn warms each austral spring, the surface elevation along our traverse should decrease. Since the UAF lidar data and ground-based GPS data were collected within a month, we expect firn compaction to have a negligible effect on our results. Conversely, the ~2 month seasonal lag between the

ATM and GPS data collection means that we may be sensitive to the seasonality of firn compaction rate, as well as any secular trend over the 4 year interval between these data sets."

(2) The authors mentioned that there were anomalous elevations at the edge of the UAF lidar swath. This raised a red flag for me: Is there a possible connection between scan angle and elevation accuracy for the airborne lidar systems (even in campaigns where elevations did not have an across-track tilt)? The authors could include elevation differences between GPS and airborne lidar as a function of scan angle, especially since their data will likely capture a fairly wide range of scan angles. This analysis will tell us if we should only use airborne lidar data from a particular range of scan angles when comparing with ICESat-2 altimetry.
AUTHORS: We do not think that scan angle is the root cause of this artifact. We note that the trough is isolated to one side (in our case the southern side) of the transect. Our expectation is that if the artifact was associated with scan angle, some matching form of the trough would also appear on the other (northern edge).

(3) It appears to me that the UAF lidar data from 3rd December 2017 were of lower quality that those from 30th November 2017 (Table 1, 2, and Figure 6 Panel C). However, biases appear to be within 1σ uncertainties, so perhaps this difference isn't significant enough to require further discussion.
AUTHORS: Our experience is that mature airborne lidar accuracies and precisions are generally under 10 cm ± 15 cm. Therefore, because data from both flights fell within this general rule of thumb, we put no stock in one having a lower bias; we do not think that this suggests that that flight is significantly better than the other. The reviewer makes a great point; and there is language (some new language) that touches on this in the text (e.g., first few lines of section 4.3, which we have augmented):
"Both altimeters had elevation biases less than 10 cm and surface measurement precisions less than 15 cm; we note that these values are similar to those in Brunt et al. (2017), which is a similar study in a similar geophysical setting."

Other minor comments:

1) Page 1, Line 20: change "set to launch" to "launched". Yay!
AUTHORS: Done! Yippee!!

2) Section 3.1: What cut-off angle was used in the processing?
AUTHORS: Good addition. We added the following text:
"We used a GPS satellite elevation mask, or a cut-off angle, of 7.5 degrees, to minimize the effects of multipath."

3) It would be nice to have larger font sizes in Figures 6 and 7.
AUTHORS: This was done in conjunction with other edits suggested by Reviewer 2.

Discussion: I think a significant number of issues with using GPS data to validate airborne and satellite altimetry could potentially be mitigated in the future with the use of a terrestrial laser scanner (TLS) mounted near the GPS antenna on the PistenBullys.
AUTHORS: We plan on attempting to make a change in the coming season that starts to get at the reviewer's point. We will try to integrate a downward looking laser, next to the GPS antenna, off of the side of the rear of the sled. This should give us an 'along-track' profile of surface roughness. While this is not provide the 3D sense of roughness around our GPS survey data (that the TLS would provide), it may help beat down the error term associated with track depth.

Thank you again,
Brunt, Neumann, Larsen