We thank all three reviewers for their constructive comments and suggestions. This manuscript will be much improved by their input. We have made changes to our manuscript. In the following responses, we use "**bold**" text for reviewer's comments, "non-bold" text for our responses, and "*italic*" for changed text in the manuscript.

Referee #2

This manuscript presents the results of a validation of ICESat-2 laser altimetry by kinematic GNSS in East Antarctica. The results of this validation highlight the high quality of both datasets, the satellite data but also the in situ validation data. Together with the validation along the 88S traverse by Brunt et al., these results provide important insights into the characteristics of the ICESat-2 mission.

The manuscript is well written, the applied methods are appropriate and the results are nicely illustrated. Alone in the structure of the manuscript I would suggest a few changes. It is sometimes difficult to follow a specific method as each section jumps from one dataset to the others. The "Data" section briefly describes the ICESat-2 data and all the measurements performed during the campaign. Then under "Methods" you describe in detail the GNSS-processing, the height reduction from the antenna to the snow surface, the ICESat-GNSS comparison and the validation with the other measurements. I would suggest to make separate subsections for GNSS, the CCRs, the RTSs and the UAV-DEM and ICESat-2 under "Data" and describe all the details in obtaining the each of these datasets there. Therefore, you could simply move the respective paragraphs from Methods to Data. Then, the "Methods" section could concentrate on the details of the validation between each dataset and ICESat-2.

(**Data** and **Method** sections) We restructured the **Data** and **Method** sections as suggested. We separated the **Data** section into subsections of **ICESat-2 data, GNSS data, CCR data, RTS data** and **UAV data**. We also moved the data acquisition related paragraphs from the **Method** section to the **Data** section, as suggested.

This little change in the structure would also allow to avoid confusion between the different types of GNSS-processing (which are otherwise easy to be mixed up). If I understand you right, you have 2 types of GNSS observations: a) the GNSS-base stations with permanent observations for ~3 days (more at Zhongshan) and b) GNSS rovers on the PistenBullys, the CCR stakes and for the ground control points of the RTS and the UAV-DEM. For the processing of the base stations (a) you use PPP. The coordinates of the rovers (b) then are obtained as differential kinematic coordinates with respect to these base stations. These coordinates are obtained in post processing (if I understand it right), so I wouldn't call that real-time kinematic (RTK). I have several remarks to that GNSS-processing:

After restructuring we put all GNSS data processing techniques into one place under **3.1 GNSS data processing** in the **Method** section where we introduced:
a) PPP: post processing for base stations;
b) PPK: post processing for receivers on PistenBully along the CHINARE route; it was called RTK in the last version, now it is corrected to PPK; and

c) RTK: real-time positioning of CCRs, RTSs and ground control points for UAV - DEM near Zhongshan Station.

1. Similar studies used reference stations at the coast (Schröder et al. 2017, doi: 10.5194/tc-11-1111-2017) or directly processed the rover GNSS-data using PPP (Kohler et al. 2013, doi: 10.1109/TGRS.2012.2207963 or Brunt et al. 2019, doi:10.5194/tc-13-579-2019). Your processing software (RTKLIB) seems to support these types processing modes. Did you try to process your rovers this way? This would overcome the limited availability of your base station data and even provide useful results for your GNSS-measurements near Taishan.

As described in the paper, PPP worked well for post processing using base stations where observations lasted up to ~3 days at each station. And PPK worked well for post processing at positions of the snowcat in motion.

We did try to process our snowcat GNSS data using the PPP technique, in addition to the PPK results. Since the quality of the PPP results are lower (see following table) the number of calculated crossovers is reduced from 26 to 21. So the PPP – PPK comparison is based on 21 crossovers. The accuracy estimated from the elevation differences at the crossovers are

Solution	**Inbound**		**Outbound**	
	Ave. (cm)	**Std. (cm)**	**Ave. (cm)**	**Std. (cm)**
PPK	0.2	5.5	-3.4	8.7
PPP	-3.9	13.4	-16.2	18.4

PPK performed better than PPP in our case. Thus, we used the PPK results. Thanks for this question regarding overcoming the limited availability of base stations in the AIS environment. We will take this suggestion into account if we would conduct another campaign in the future.

Due to logistic difficulties, GNSS-measurements near Taishan Station were carried out using the single-point positioning technique. The GNSS function is integrated into a field surveying pad system. The data are not appropriate to support a PPP solution. Similarly, we will also look into a potential PPP application for these single point situations in the future.

2. It is quite usual that internal accuracy values reported by the GNSS-software are way to optimistic. You results for internal crossovers in your GNSS-profiles demonstrate that nicely. However, in the vertical GNSS accuracy at the CCR of the ground control points, you simply state values of ~0.3-0.4 cm without any information about their origin. Is that the accuracy reported by the software? For how long was each point observed? This accuracy is remarkably high for a short observation. Do you have any evidence, that this value is realistic?

In the **GNSS data** subsection we added text to address observation time "*……10 CCRs, 137 randomly distributed GNSS points on the RTS, and 3 GCPs for UAV geometric control near Zhongshan Station were surveyed using the RTK technique. The observation time at each point was about 4 to 5 seconds ……*"

In the **Method** section under **3.1 GNSS data processing** we added a paragraph to address where the internal precisions are from:

"*The RTK positioning technique is applied to estimate positions of the CCRs and GNSS points on the RTS sheets near Zhongshan Station. We used a known GNSS control point at Zhongshan Station as a reference point for RTK. The CCR and RTS checkpoint positions were estimated in real-time by the GNSS receiver's onboard software (https://www.chcnav.com/uploads/i70_DS_EN.pdf). Additionally, the UAV - DEM reconstruction was geometrically controlled using 3 GNSS GCPs. The positions were estimated by the RTK technique implemented in the UAV package (https://www.dji.com/hk-en/phantom-4-rtk?site=brandsite&from=nav). The accuracy of the RTK positions is estimated based on internal precisions given by the applied GNSS systems and the accuracy of the reference point*."

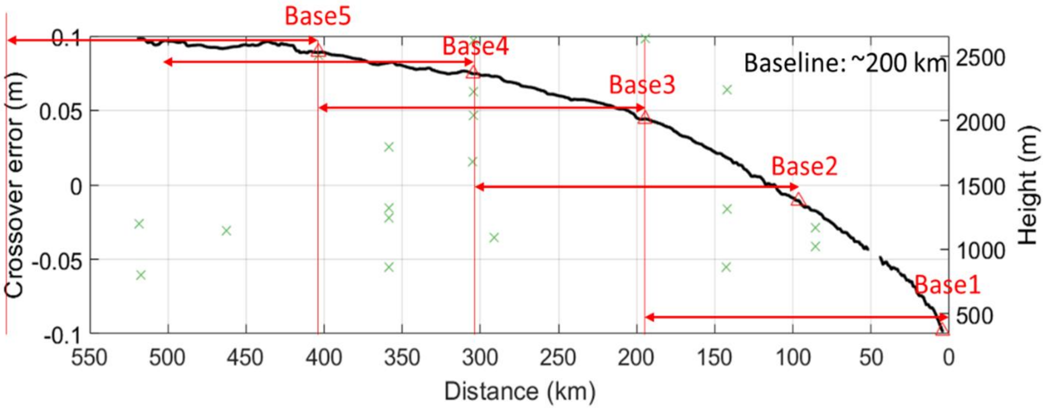
The GNSS system reported internal precisions of individual CCRs, from which an overall RTK internal precision (optimistic) is estimated as 0.3 cm (horizontal) and 0.4 cm (vertical), respectively. Sorry, we did not account for the GNSS reference point error in the last version of the paper. The GNSS reference point was surveyed four times (~7-10 hours each time). The elevations are (-75.9516 m, -75.9617 m, -75.9553 m, -75.9286 m) with 1σ of 1.2 cm. Thus, the overall accuracy of CCRs is 1.3 cm.

In the **Results** section we revised the text to give external accuracy "*……Around 6 h before the ICESat-2 pass, the CCRs were deployed and RTK GNSS was surveyed. Based on the internal precisions of individual CCRs given by the GNSS system and accuracy of the known GNSS reference point, the elevation accuracy of ten CCRs is 1.3 cm…...*"

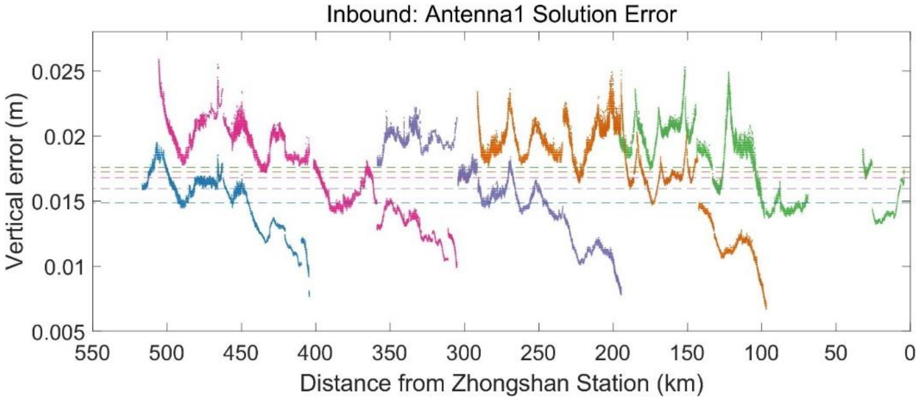
3. (this comment refers to l.240) I appreciate that you checked the accuracy of the GNSS-profiles using crossover differences at intersections. However, I would suggest a few more analyses:

- If I understand you correctly, you separate your total profile into shorter sections according to the distance to the base stations and post-process each section with this base station as reference. It would be very interesting to have some overlap in the processing of these the sections. Hence, at the transition from one section to the next one, you could process some measurements with both base station and compare the results.

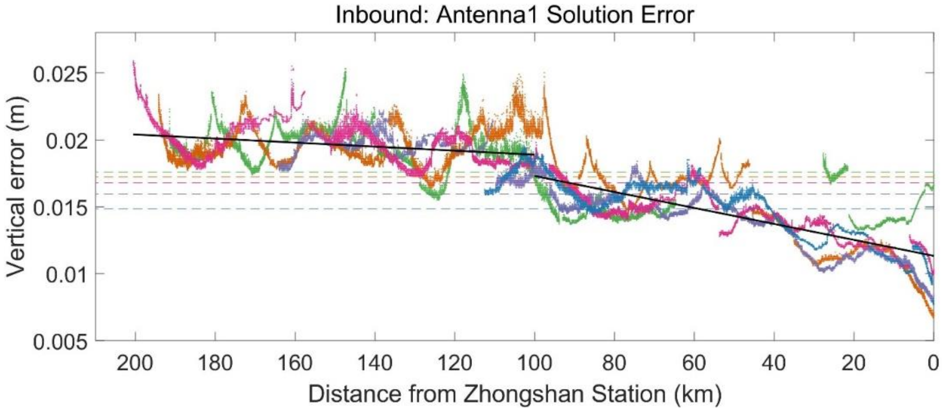
Thank you for this interesting suggestion. Constrained by our familiarity to the software system, available time, and level of our understanding of your suggestion, we may not have implemented it in the exact way you want. However, we may have got it close. From each base station we extended each segment from ~100 km to ~200 km (red arrowed lines in the following figure), along which we used PPK technique to solve for rover positions (use only one base station, instead of two).



We plot the internal precision profiles of each ~200 km inbound segments in the following figure (colors do not have a specific meaning). They show that there seems to have a linear trend of decreased precision as distance from a base station increases. The extended segments from ~100 km to ~200 km generally have greater vertical errors than the first half. Overall, the profiles of the flat inland interior (beyond ~100 km from Zhongshan Station) showed better performance (particularly in their first ~100 km).



We then moved each profile to start from the same origin in the following figure to examine the internal precision trend vs. distance controlled by a single base station. The linear trend is clear. Thus, we use a linear function $y = k \cdot x + b$ to model the relationship between the internal precision y and distance from coast x. We have three cases: Traverse (0 km – 200 km), Segment 1 (0 km – 100 km), and Segment 2 (100 km – 200 km). The average ± standard deviation of the internal precisions for the three cases are 1.6 ± 0.3 cm, 1.4 ± 0.2 cm, and 1.9 ± 0.2 cm, respectively.



Within the first ~100 km (Segment 1) errors increase at a steeper rate of 0.6 cm per 100 km and reach up to

~1.7 cm; this may be mainly attributed to the effect on the vehicle by rugged terrain and topographic change from coast to inland. The rate of the second half (Segment 2) is lower at 0.2 cm per 100 km where the ice surface is flat.

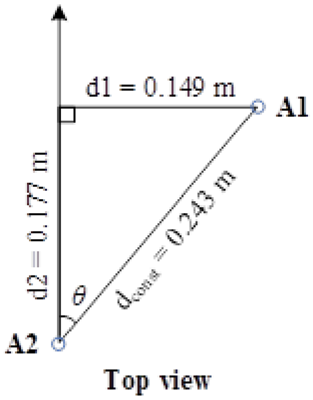
Segment	k (cm / 100 km)	b (cm)
Traverse	0.5	1.2
Segment 1	0.6	1.1
Segment 2	0.2	1.7

On the other hand, we also use elevation differences at crossovers to assess the accuracy before and after the extension of the controlled distance of the GNSS base stations. The extension of the segments from ~100 km to ~200 km resulted in a non-significant uncertainty increase from 0.6 ± 5.4 cm to 1.2 ± 5.9 cm.

In summary, based on the above analysis the extension of the segments from ~100 km to ~200 km does not cause a significant increase of the internal and external uncertainties. Since this work is still being optimized and the detailed analysis will be presented in a dissertation, we hope that you would agree that we will not put the detailed result in this paper. However, it should guide us in designing our future work in potentially another expedition.

- Moreover, referring to l.139 you have two antennas. Are they installed on the same PistenBully? If yes, are there any systematic offsets between them and what is the noise? If no, did you check crossovers between the two PistenBullys?

Antenna 1 and Antenna 2 are installed on the roof of Pistenbully Polar300 (one vehicle) with a fixed offset (24.3 cm as measured by a steel tape).



The configuration was designed to have a check of GNSS observations for the hard baseline. However, as stated now in the **2.2 GNSS Data** subsection: "…… *Antenna 2 served an ice penetrating radar equipment during the inbound trip. Due to inter-equipment interferences and incidental battery problems, the GNSS rover surveying was carried out by a combination of two receivers……*", the data rate for Antenna 2 was set as 10 Hz to match that of the radar device, compared 1 Hz for Antenna 1. That caused drifted positions of Antenna 2 with an uncertainty of -8.3 ± 10.0 cm as assessed with crossovers, compared to 0.2 ± 5.5 cm for Antenna 1.

We consulted with the venders and they confirmed the potential interferences between the devices, which they have also encountered before with the same models of devices. The batteries had a few incidental problems. Thus, we ended up with covering the entire traverse with only one antenna at any time (inbound with Antenna 1; first ~80 km of outbound with Antenna 1, rest of outbound with Antenna 2 when the radar device was off), and being unable to use two antennas to do a reasonable check of the hard baseline.

Besides these major point and remarks, I have following detailed comments:

l.21: "...which is important for overcoming the uncertainties in the estimation of mass balance in East Antarctica" is a very general statement. The most important topics in East Antarctic mass balance are probably eliminating mission biases and the conversion from volume to mass. This validation contributes only to the first point.

The sentence is revised: "…… *which is important for eliminating mission biases by overcoming the uncertainties in the estimation of mass balance in East Antarctica*……"

l.30: remove "As"

Removed.

l.55-74: I suggest to add brief motivations for each of the methods of validation (kinematic profiles, CCRs, RTSs, UAVs). It would be useful for the reader to know the benefit of each of the methods from the very beginning of the paper.

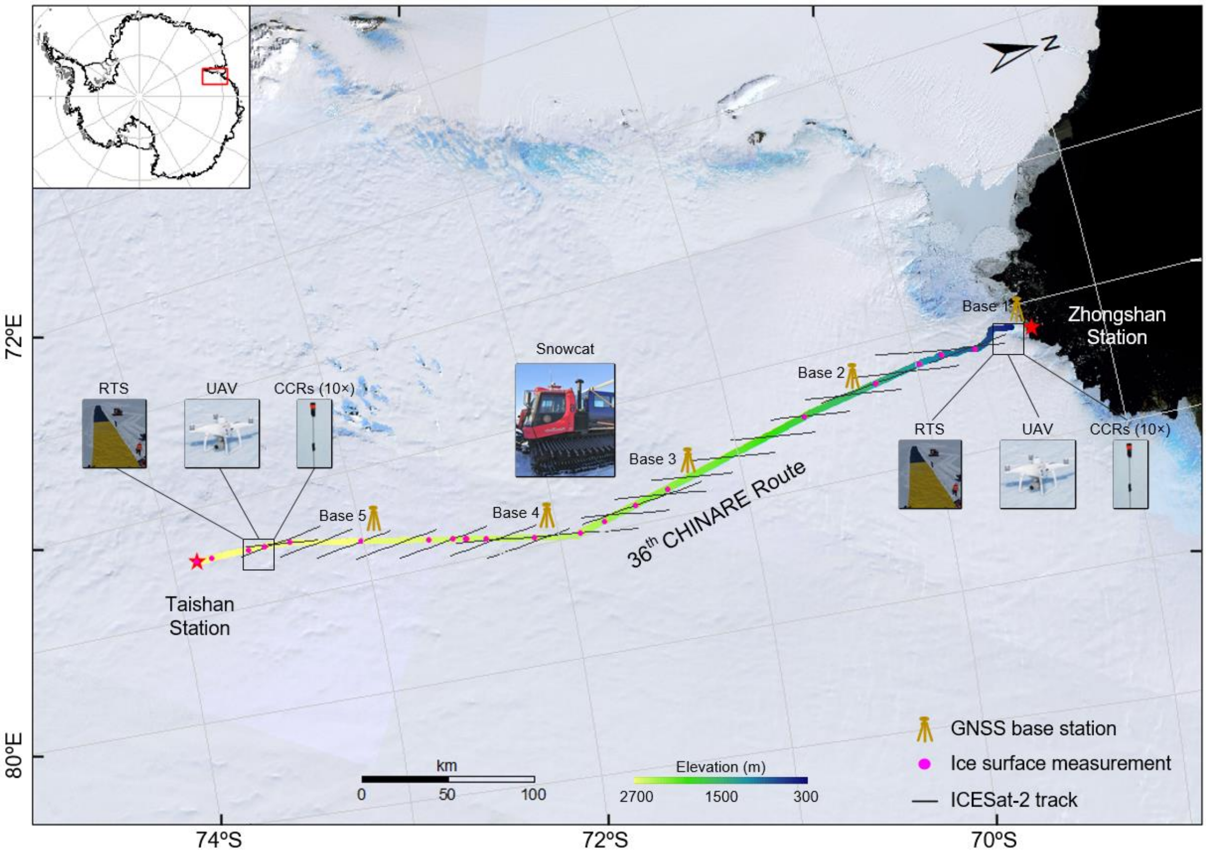
We changed it to a separate paragraph to serve this purpose:
"*In order to validate the ATL03 and ATL06 data along the CHINARE route from the coastal Zhongshan Station to the inland Taishan Station, two roving GNSS receivers of CHC i70 from CHC Navigation Technology LTD (http://www.huace.cn/product/product_show/291) were installed on roof of a snowcat, Pisten Bully Polar 300, to measure ice surface elevations using the post processed kinematic (PPK) positioning technique. Supported by the precise point positioning (PPP) technique, five GNSS base stations with CHC i70 receivers were deployed every ~100 km along the traverse to enable the PPK positioning of the vehicle. Two line arrays of ten upward-looking CCRs (optical prisms) with known elevations were deployed at sites near Zhongshan Station and Taishan Station, respectively, to reflect photons for the verification of individual photons. We used one rectangular (5 m × 150 m) RTS for each site to investigate the reflectivity and elevation accuracy of photons reflected from selected RTS coatings. Finally, two UAVs, DJI Phantom 4 (https://www.dji.com/hk-en/phantom-4-rtk?site=brandsite&from=nav), were used to acquire images for the generation of digital elevation models (DEMs) for an areal assessment of ICESat-2 elevation accuracy. The real-time kinematic (RTK) positioning technique was applied to provide horizontal and vertical positions of the CCRs, GNSS points on RTSs, and control points for UAV - DEM reconstruction*......"

l.64: Later you describe that you obtain your coordinates in post-processing. So the survey is not real-time kinematic (RTK). This totally makes sense (as post-processing is much more precise) but please be also precise in describing your method.

They are corrected. We have changed "RTK" to "PPK" in related places in **Introduction** and elsewhere.

Fig.1: In this context, no measurements have been conducted at Great Wall Station. It is fair to mention the original plans in the text but I suggest to remove the inset from Fig.1.

(Figure 1) The Great Wall Station is now removed in the inset.



l.95: How is this accuracy obtained? You describe the GNSS-processing and the validation methods later, so such accuracy measures should appear later, when the reader knows how they were obtained. Furthermore, is this really RTK (see comment on l.64)?

As you suggested, the **Data** section is restructured. This part is now under GNSS data subsection.

(Also see above response "Line 64") It should be "PPK". We have changed "RTK" to "PPK".

Now we clarify the positioning techniques of PPP, PPK and RTK in the **Introduction** section.

The description of the PPK accuracy is given in the **Method** section:
"*The internal precisions of the estimated positions from the PPP and PPK processing are given by the software systems. Furthermore, we used the accuracy computed from elevation differences at crossovers where the GNSS surveyed track intersected itself, as shown in Fig. 2c. These crossovers are the intersections of tracks by the snowcat during instrument installations, observations, and overnight breaks. Within a neighborhood of the intersection, we fit two lines to compute the crossover location and elevation difference (Kohler et al., 2013).*"

We moved the estimated PPK accuracy from the **Data** section to the **Results** section:
"*…… which have an average internal elevation precision of 1.6 ± 0.6 cm given by the software system. Finally, the elevation accuracy of the GNSS traverse was assessed as 0.3 ± 5.8 cm by using 26 crossovers of the traverse itself (Fig. 2c) where the GNSS surveyed elevations from two intersecting traverse segments were*

compared.”

l.100: See comments on l.95

The general description of the RTK accuracy is given in the **Method** section:
“*…… The accuracy of the RTK positions is estimated based on internal precisions given by the applied GNSS receivers and the accuracy of the reference point.*”

The estimated RTK accuracy for CCRs and RTSs are now given in the **Results** section: “*…… Based on the internal precisions of individual CCRs given by the GNSS system and accuracy of the known GNSS reference point, the elevation accuracy of ten CCRs is 1.3 cm.…..*”

l.113: Start a new sentence after “532 nm”.

It is changed accordingly.

l.134: What is the reference frame and the ellipsoid for the GNSS coordinates and the ICESat data? And how about the permanent tide? Many altimetry measurements refer to the Topex ellipsoid and are in the ‘mean tide’ system while GNSS data generally refer to WGS84 and are ‘tide-free’. Please give some details here to show that the different data are comparable.

We added a paragraph in the **Method** section to clarify the reference frames:
“*In the ICESat-2 products geographic coordinates (latitude and longitude) are defined based on the WGS84 ellipsoid and heights are referenced to the ITRF2014 frame (Brunt et al., 2019; Neumann et al., 2019); corrections for solid earth tides, ocean loading, solid earth pole tide, ocean pole tide and others are applied to the ATL03 data (Neumann et al., 2019). On the other hand, the processed GNSS data are also referenced based on the WGS84 ellipsoid (Schröder et al., 2017); the ITRF2014 reference frame is used in the GFZ precise ephemeris and precise orbit products which is input to the RTKLIB and MUSIP post processing systems; furthermore, the geophysical corrections for the above tides are applied (Petit and Luzum, 2010). Thus, the reduced ice surface heights are “tide-free” and the permanent crustal deformation is removed (Schröder et al., 2017; Brunt et al., 2021).*”

The cited papers are added to **References**:

l.139: Were both antennas mounted on the roof of the same PistenBully?

Yes, please see the above response to “**3. / – Moreover ……**”

l.179: The description of the interpolation of h2 fits much better in section 3.1.2.

As suggested, we moved the sentence to section 3.2.1.

l.210: I would have expect that the CCR looks like a unique point reflector. Why does it show up as a curve with slightly lower elevations on the sides?

The photon streak length is a function of the CCR diameter, among other parameters. The mission team used an 8 mm diameter CCR that resulted in a flat streak of ~11 m (Magruder et al., 2020). We used a 6 cm diameter

CCR and obtained a ~38 m near Taishan Station (and ~34 m near Zhongshan Station) curved streak.

The following is extracted from the responses to Reviewer #1's comments.
The central disc of our 6 cm aperture CCRs has a diameter of ~10.82 m in comparison to ~40.57 m of the ~8 mm diameter CCRs in Magruder et al. (2020), according to Chang et al. (1971). Here is the Fraunhofer diffraction pattern of our CCR:

$$D = \frac{1.22\times\lambda\times H}{d}\times 2.$$

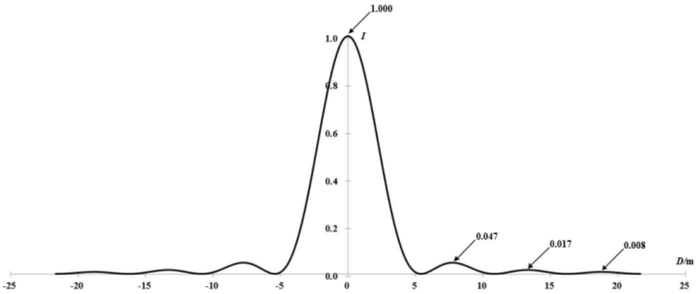


Figure of the Fraunhofer diffraction pattern of the 6 cm diameter CCR (λ=532 nm, H=500 km, and d=6 cm)

With the larger CCR aperture, in addition to the smaller central disc diameter, the total signal level (for both central disc and lobes) is also increased. Thus, it makes signals from the lobes of the Fraunhofer diffraction pattern also detected. In the following figure, at an approaching position ($t_{\text{Approaching}}$) ATLAS received and accepted signals from lobes of both the nadir CCR #6 (red signal curve) and the neighboring CCR #7 (light blue signal curve), both at a lower signal level; this resulted in the reflected photons of higher elevations (green dots) of CCR #6 and those of lower elevations (green dots) of CCR #7. However, at the nadir CCR position (t_{Nadir}) ATLAS received and accepted high level signals from the central disc of the nadir CCR #6 (higher elevation), but may have rejected the lower level signals from the neighboring CCR #7 (lower elevation) because of the much increased ratio between the signals. This allows us to determine the window size, ~9 m gap of the lower streak, to select photons of the nadir CCR (including those inside the central disc) for CCR elevation estimation.

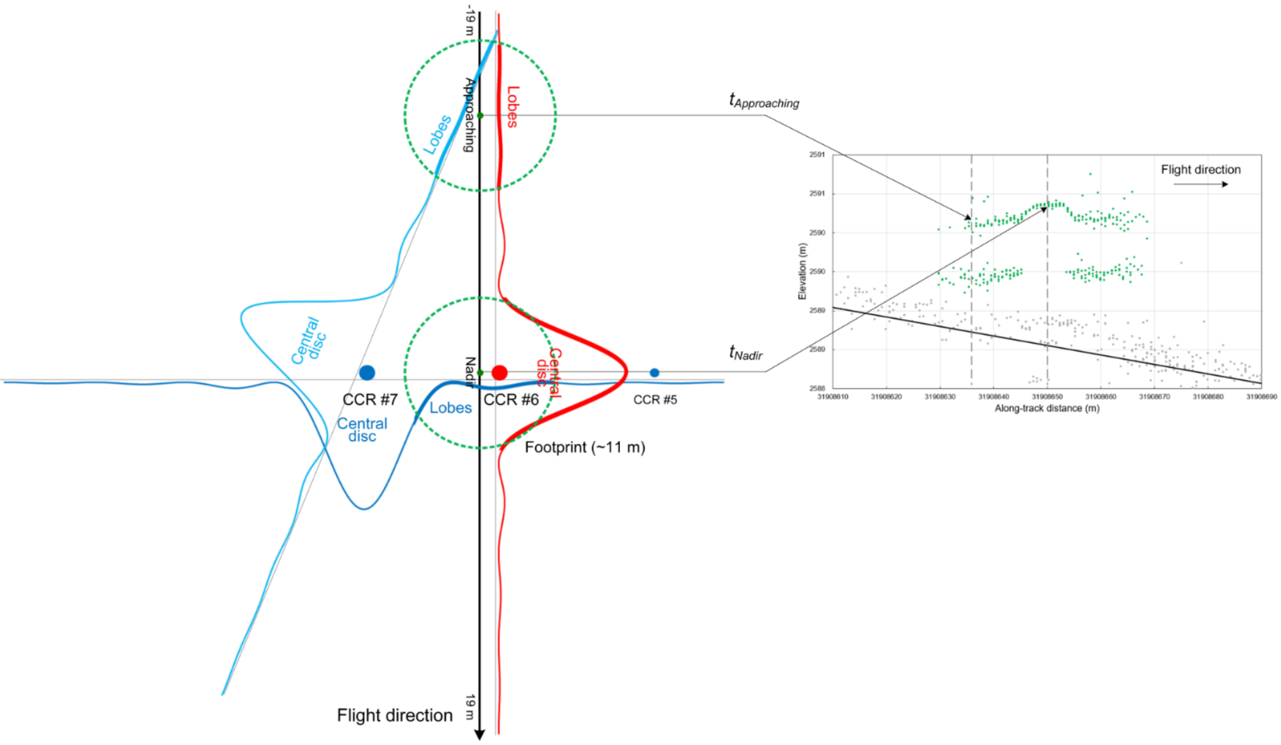


Figure for CCR signal analysis at the site near Taishan Station

We added a paragraph in the **Discussions** section to explain the impact:

"*The use of the readily available CCRs of 6 cm diameter for the 532 nm wave length of ATLAS, which is larger than 8 mm of the CCRs used in Magruder et al. (2020), is subject to velocity aberration caused by a decreased central disc and receiving signals from the outer lobes of the Fraunhofer diffraction pattern (Chang et al., 1971; Born et al., 1999; Sun et al., 2019; Magruder et al., 2020). In addition, the larger aperture of the CCR resulted in a higher level of the total signals received by ATLAS so that signals from both the smaller central disc and outer lobes are detected and used to estimate elevations in ATL03 data. This may have attributed to the creation of the long along-track streaks of ~35 m (Fig. 6a) and ~38 m (Fig. 6c) in comparison to those of ~11 m in Magruder et al. (2020). Thus, photons reflected from the lower neighboring CCR(s) in the cross-track direction (Fig. 6c) were detected for the same reason. Similarly, the one-layer photon streak (green dots in Fig. 6a) may include those reflected from one or both neighboring CCRs because the elevations of all three CCRs (#4, #5 and #6) are within a 15 cm range (Table A1) due to local ice surface topography and logistic constraints, although the poles were manufactured in different lengths. On the other hand, the received signals in the central disc are generally higher (about 84% of the total energy) than in the outer lobes given atmospheric scattering and other optical losses (Magruder et al., 2020). Correspondingly, we observe that within the window of the nadir CCR (red rectangle in Fig. 6c) the photons are densely aligned along a curve. The same curve trend appears to continue towards both ends, but diverged by potentially blended signals reflected from neighboring CCRs (Figs. 6a and 6c). Therefore, by selecting photons inside the central window of the CCR streak it ensures that high quality photons in the central disc of the Fraunhofer diffraction pattern be used to estimate the elevation of the representative photon of the nadir CCR through the fitting curve. The result is also validated by the nadir CCR position surveyed by using the high-precision GNSS RTK technique.*"

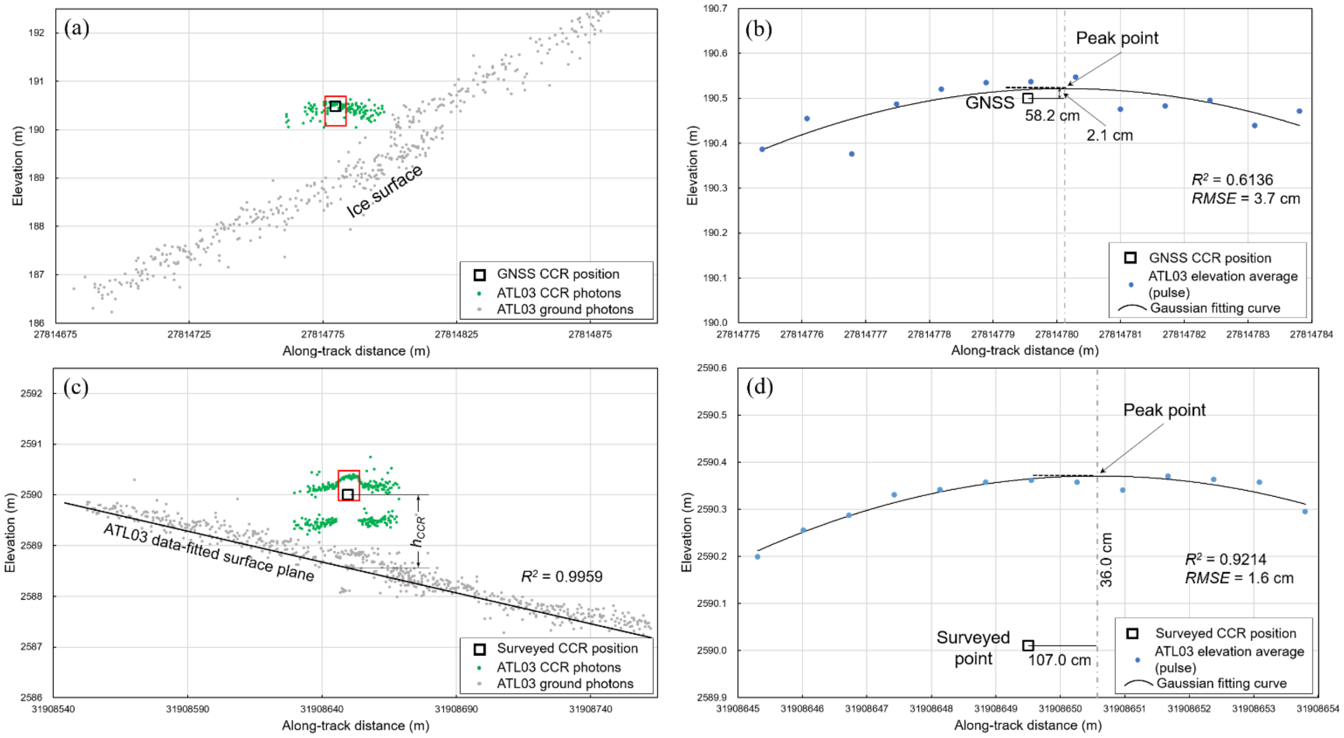


Figure 6. (a) CCR experiment near Zhongshan Station: returned photons (ATL03), GNSS-surveyed CCR position, and ice surface photons (ATL03); (b) elevations averaged in each pulse in the red rectangle in (a) to compare with the GNSS-surveyed position; (c) CCR experiment near Taishan Station: returned photons (ATL03), steel tape-surveyed CCR position, and h_{CCR} - height between CCR center and ice surface (ATL03); and (d) elevations averaged in each pulse in the red rectangle in (c) to compare with the steel tape-surveyed position.

l.226: How was this accuracy obtained? From l.230, I guess that the accuracy of the UAV-DEM is just several meters and you need several ground control points for a more precise absolute orientation. Is this correct? Could you explain this a bit more detailed (or give some references)?

We revised the paragraph in the **Method** section to clarify the photogrammetric process. We also gave a reference for photogrammetric orientations.

"In the mapping area a set of ground control points (GCPs) are surveyed using the RTK positioning technique. They are further used to perform a photogrammetric absolute orientation (McGlone, 2013). The 3D surface points are reconstructed from the UAV images by using the structure-from-motion multi-view stereo (SfM-MVS) algorithm (James and Robson, 2012; Turner et al., 2014) implemented in the Pix4Dmapper software (version 4.5.6, https://support.pix4d.com/hc/en-us/categories/360001503192-Pix4Dmapper). As the result, a UAV-DEM and an orthophoto at a centimeter level accuracy (both horizontal and vertical) are generated. Thereafter, we evaluate the elevation differences ΔH *between the elevations of the ICESat-2 ATL06 ice surface points (*$H_{Ice\ surface}$*) and the corresponding elevations of the UAV-DEM (*H_{UAV_DEM}*):*

$$\Delta H = H_{Ice\ surface} - H_{UAV\_DEM}. \quad (3)"$$

In the **Results** section we present the estimated accuracy:

"……*A GCP-controlled photogrammetric processing of the UAV images was successfully performed with an internal precision given by the software as 2.1 cm (horizontal) and 2.8 cm (vertical), respectively. The* *elevation accuracy of the generated UAV-DEM was then evaluated as 0.2 ± 6.3 cm using 167 GNSS RTK points. On the other hand, the 1L (weak beam) and 1R (strong beam) tracks have 48 ATL06 ice surface points in the DEM area, among which three were affected by the photons from the CCR and excluded from the validation. The elevation differences between the ATL06 ice surface points and the UAV-DEM were computed and resulted in an estimated ICESat-2 ice surface elevation uncertainty of 1.1 ± 4.9 cm.*"

l.264: With regard to the variation between the different ground tracks, the lack of GT2-results for ATL03 (due to the very strict exclusion conditions) and the precision of ~9 cm I would be careful concluding that the ATL06 bias is smaller than the ATL03 bias. I don't think that the differences are significant.

We agree with you and changed the text accordingly. Please see the response to **Comment l.333** (last page).

l.266ff: You state that the standard deviation of the h2 measurements is ~3cm and this variation is mainly attributed to the microtopography and firn density changes. However, the effect of microtopography will apply immediately when you move a few meters. Systematic variations in density should be largely accounted for by the IDW interpolation. So in my opinion, these effects alone cannot be responsible for these larger differences when using the full profile. Using only data in a 2 km vicinity of the h2 measurements reduces the usable crossovers dramatically. So, before reducing the amount of data so drastically, I suggest to do a few more analysis on these larger difference. Did you do any outlier checks before calculating these standard deviations? A few outliers can have a large effect on stddev. Or is there a specific spatial pattern (maybe in the region around 71S, which is the largest gap between h2-measurements)?

That is a good suggestion, thanks. The following plot of ATL06 – GNSS elevation differences (red and blue dots) does show that there is a segment of traverse in the red rectangle where outliers occurred. So, we checked the field measurement notes that show that the closest three h_2 measurements (diamonds on the bottom) were erroneously measured to the bottom of wheel-chain prints, while all other h_2 values were measured to the ice surface. They were the first three ice surface measurements of the traverse.

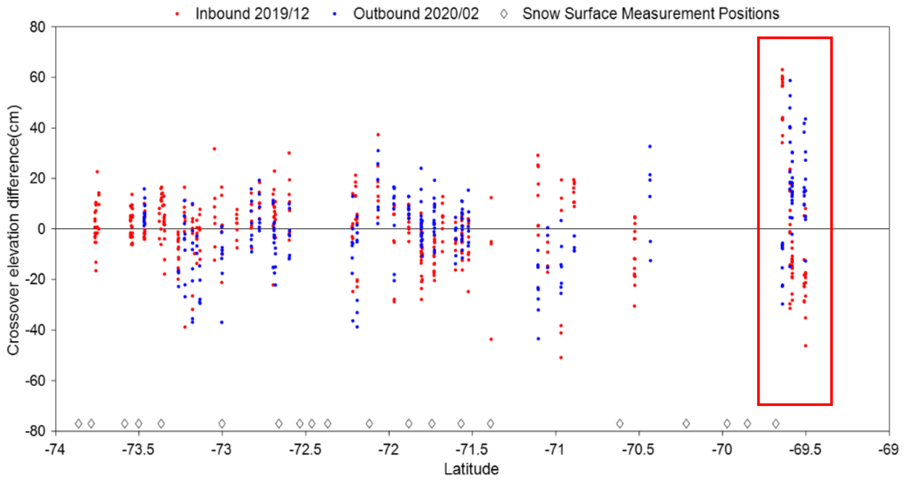


Figure of elevation differences at crossovers between ICESat-2 (ATL06) tracks and the GNSS traverse with red dots for inbound and blue dots for outbound pairs, and diamonds for h_2 measurement locations.

We performed the following experiments to test the impact of the outliers: a) including just each one of the three h_2 outliers (along with rest), and b) including all three h_2 outliers (along with rest; this had the lower precision result in the appendix of the previous manuscript). No distance limit to h_2 locations is applied. All results of Experiment a) are similar to that of Case b). Thus, we had to give up all three h_2 measurements (outliers), correspondingly also the ATL06 – GNSS pairs in the red rectangle in the above figure since they are now greater than ~65 km away from the closet h_2 measurement.

In the further analysis, we have two sets of results: Table 2. Results with a limit of ~5 km distance from h_2 measurements, and Table B1. Results with all intersections along the traverse without limit of distance.

Table 2 Assessment of ICESat-2 ATL06 ice surface points and ATL03 photons using the GNSS PPK technique with direct ice surface measurements (h_2) within ~5 km along the 36th CHINARE traverse. Bias and precision were estimated from their elevation differences using N ice surface points or photons. The difference is calculated as ICESat-2 elevation minus GNSS elevation.

Ground Track	ATL06 Bias ± Precision(cm)	ATL03 Bias ± Precision(cm)
GT1L	+2.7 ± 9.6 (N = 64)	+5.9 ± 5.9 (N = 1518)
GT1R	+3.0 ± 7.3 (N = 62)	+1.7 ± 6.7 (N = 2608)
GT2L	+0.7 ± 7.9 (N = 48)	-0.5 ± 6.7 (N = 862)
GT2R	- 2.3 ± 12.0 (N = 42)	+5.8 ± 14.0 (N = 1356)
GT3L	+1.3 ± 8.4 (N = 33)	+4.2 ± 7.7 (N = 800)
GT3R	- 0.7 ± 8.7 (N = 36)	+4.6 ± 10.9 (N = 2695)
ALL	+1.5 ± 9.1 (N = 285)	+4.3 ± 8.5 (N = 9839)

Table B1 Assessment of ICESat-2 ATL06 ice surface points and ATL03 photons using the GNSS PPK technique at all intersections along the GNSS traverse. Bias and precision were estimated from their elevation differences using N ice surface points or photons. The difference is calculated as ICESat-2 elevation minus GNSS elevation.

Ground Track	ATL06 Bias ± Precision(cm)	ATL03 Bias ± Precision(cm)
GT1L	+2.9 ± 12.0 (N = 111)	+6.4 ± 12.6 (N = 2055)
GT1R	+2.5 ± 12.7 (N = 116)	+4.5 ± 12.5 (N = 4542)
GT2L	-0.3 ± 12.4 (N = 101)	+0.2 ± 7.3 (N = 1245)
GT2R	-2.8 ± 12.4 (N = 106)	+0.5 ± 11.8 (N = 4272)
GT3L	-1.1 ± 12.6 (N = 88)	+2.8 ± 10.6 (N = 1233)
GT3R	-2.6 ± 13.0 (N = 99)	+1.7 ± 12.1 (N = 4490)
ALL	+0.5 ± 12.7 (N = 621)	+3.4 ± 11.5 (N = 17839)

It is observed that by increasing the distance (to h_2 measurements) from ~2 km to ~5 km, the number of observation samples N in Table 2 increased significantly, and the data gaps in tracks of GT2L and GT2R are also filled; the accuracy is higher than that in Table B1 for the full profile. On the other hand, although by eliminating the h_2 outliers the standard deviations for the full profile decreased from (15.7 cm, 14.2 cm) in the last manuscript to (12.7 cm, 11.5 cm) in Table B1, they are still relatively larger than (9.1 cm, 8.5 cm) in Table 2 for the ~5 km distance constraint. Therefore, we used the result in Table 2 in this revised version, and put Table B1 in the Appendix.

Accordingly, the text in the **Results** section is changed:

"*There were 20 locations along the traverse (Fig. 1) where* h_2 *was measured (Fig. 4a). Three outliers at the beginning of the traverse were eliminated. Along with other two bore sight parameters,* h_0 *and* h_1*, these direct measurements were used to derive the ice surface elevations* $H_{Ice\,surface}$ *from the roof-mounted kinematic GNSS observations* H_{GNSS}*. The average of the measured* h_2 *is 94.0 cm with a standard deviation of 2.8 cm. This variation is mainly attributed to the microtopography and firn density changes at different locations along the 520 km traverse from the coast to the highland interior. Out of the 134 intersections between the GNSS traverse and ICESat-2 tracks, we selected 60 intersections that are within 5 km of the* h_2 *direct measurement locations to enhance the comparability between the elevations observed by the ICESat-2 satellite and the kinematic GNSS receivers along the traverse. The average distance between the intersections and measurements is ~2366 m. We validated the elevations of the ICESat-2 ATL06 ice surface points and ATL03 photons using the GNSS-surveyed elevations that are summarized according to six ICESat-2 tracks separately (Table 2 above).*

Table 2.

Compared to the kinematic GNSS elevation observations, the ATL06 ice surface points have median elevation differences (bias) for the six ICESat-2 tracks ranging from -2.3 cm to 3.0 cm and precision values (1σ) ranging from 7.3 cm to 12.0 cm, resulting in an overall bias of 1.5 cm and precision of 9.1 cm. Similarly, the ATL03 photons have an overall bias of 4.3 cm and precision of 8.5 cm. No significant elevation differences were found between the tracks of the weak and strong beams. The difference between the bias of 1.5 cm for the processed ATL06 application product (L3A Land Ice Height data) and that of 4.3 cm for the unprocessed ATL03 product (L2A Global Geolocated Photon Data) is considered insignificant, taking their precision values, 9.1 cm and 8.5 cm, respectively into account.

We further extended our assessment to all intersections of the ICESat-2 tracks and the GNSS traverse without the above 5 km selection constraint. At each intersection, the h_2 *value was calculated between two measurement locations using the IDW interpolation method. As shown in Table B1, the ATL06 and ATL03 data present a bias of 0.5 cm and 3.4 cm, respectively, which are comparable to these in Table 2. However, the overall precision values of 12.7 cm (ATL06) and 11.5 cm (ATL03) in Table B1 are relatively larger than 9.1 cm (ATL06) and 8.5 cm (ATL03) in Table 2.*

l.282 Without a precise GNSS-elevation of the CCR, you are comparing ICESat-2 measurements to ICESat-2 measurements of the same orbit. So, all you can do is comparing the offset between photons from the ground and from the CCR to the measured stake height. I would suggest to do this using the ATL03 data only as you cannot be sure, which photons contributed to the mean ATL06 data point.

As suggested, we modified Figure 6 (a) and (c) (see Fig. 6 above). We recalculated the CCR elevation offset using the ATL03 data along with the stake height. The text is revised:

"*CCR #6 was found to have returned 52 photons in 13 pulses from the weak beam track (2L) within the ~9 m window in Fig. 6c. To estimate a CCR #6 elevation that is more accurate than the meter-level GNSS result, we first used the ATL03 ice surface photons (black dots in Fig. 6c) to fit a terrain surface plane with an* R^2 *of 0.9959. Using the measured CCR height* h_{CCR} *in Fig. 6c and the fitted ice surface, the improved CCR* *elevation (black square) was calculated. After that, the photons were averaged within each of pulse (pulse id from 19 to 31, blue dots in Fig. 6d) to fit another Gaussian function with an* R^2 *of 0.9214 and an RMSE of 1.6 cm. The peak position of the Gaussian function was used as the representative photon of the CCR that has an offset of 107.0 cm in the horizontal direction and 36.0 cm in the vertical direction from the estimated CCR location.*"

l.333 As discussed on l.264, I don't believe that the difference in the offset between ALT03 and ATL06 is significant.

Now we filled the data gaps of GT2L and GT2R in Table 2. The difference between the bias of 1.5 cm for ATL06 and 4.3 cm for ALT03 still exists. However, considering their standard deviations of 9.1 cm and 8.5 cm, we agree with you that this difference is not significant.

In the **Results** section we revised it to: "…… *The difference between the bias of 1.5 cm for the processed ATL06 application product (L3A Land Ice Height data) and that of 4.3 cm for the unprocessed ATL03 product (L2A Global Geolocated Photon Data) is considered insignificant, taking their precision values, 9.1 cm and 8.5 cm, respectively into account.*"