The authors would like to thank the referee for the insightful review of the manuscript and the suggestions for the revision of the manuscript. Replies to the comments by the referee are as follows. The original comments from the referee are shown in italic Arial font, and the reply to each comment is provided immediately after the comment and is in Times font.

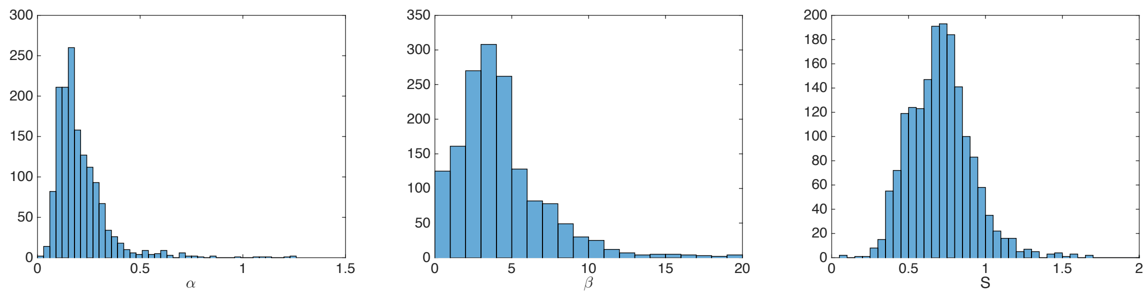
Comments A:

The major concerns I have about the presented study are stated here:

*a) The authors do not discuss or show any results at all on the advantage of their suggested approach. If the new contribution here is to combine L-band and altimeter measurements, why do you not show the difference between using this combination of TB and freeboard measurements as compared to just using freeboard measurements and the relationship that you found between snow depth (hs) and FB? This could be done by 1) finding global alpha and beta values (instead of a global s=alpha*beta) and then using these alpha and beta values and Eq. 3 to convert FB to hs and then ice thickness (hi) (using Eq. 1 afterwards) or 2) finding a good fitting formulation that relates hs and hi and then converting FB to hs and hi using Eq. 1. The correlation between FB and hs (R^2=0.67) is very similar to the correlation found between the retrieved and the OIB observed hs for using the global s values and SMOS TBs (R^2=0.64) / simulated TBs (R^2=0.65). Maybe it is essentially the correlation between FB and hs that is behind this agreement of retrieved and observed hs values? From the presented results, I cannot see whether there is any advantage in using L-band additionally...*

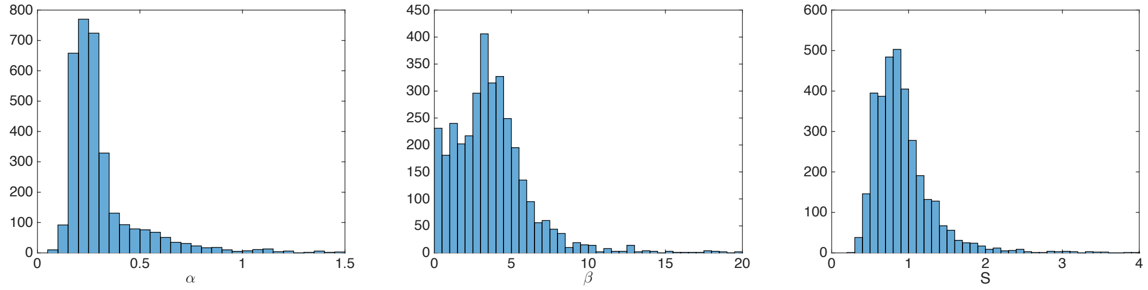
The authors would like to emphasize that the main advantage is <u>the simultaneous retrieval of BOTH ***Hi*** and ***Hs***</u>, by combining L-band passive remote sensing and laser altimetry. The covariability information that is incorporated into the retrieval does NOT specify ***Hs***. Rather, this information is only a general constraint between $\boldsymbol{FB_{snow}}$ and ***Hs***.

According to the suggestions from the referee, we further carried out "retrieval" study using only the statistics derived from OIB, and NOT using TB. Specifically, the statistics are computed for the parameters ***alpha***, ***beta***, and ***s***. Results are shown in Figure 1 and 2 (below) for FYI and MYI, respectively. Figure 3 (below) shows the retrieval results for ***Hi*** and ***Hs*** using the globally fitted ***alpha*** and ***beta*** on the exactly same spatial scale of the retrieval in the manuscript. The adopted values for these parameters are specific to FYI and MYI: for FYI, ***alpha***=0.21 and ***beta***=3.38; for MYI, ***alpha***=0.31 and ***beta***=3.07.



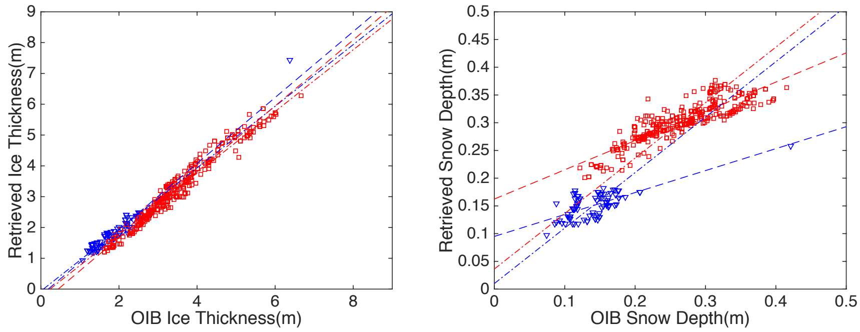
(a) ***alpha***. (b) ***beta***. (c) ***s***.

Figure 1. Distribution of alpha, beta and s for FYI. 40-meter resolution (OIB) data are used for computing the value of each parameter on the scale of 37.5 km (i.e., approximately SMOS TB resolution)



(a) ***alpha***. (b) ***beta***. (c) ***s***.

Figure 2. Distribution of alpha, beta and s for MYI. Specifications are the same as Figure 1.



(a) ***Hi***. (b) ***Hs***.

Figure 3. Retrieval of ***Hi*** and ***Hs*** using globally fitted ***alpha*** and ***beta***.

As shown in Figure 3, with the covariability information (***alpha*** and ***beta***), indeed retrieval can be carried out. The least squares fit (dashed lines) for FYI (blue) and MYI (red) for ***Hi*** yields 0.93 and 0.96 respectively for R^2, and that for ***Hs*** yields 0.51 and 0.77. However, the least squares fit (dotted-dashed lines) for ***Hs*** on FYI (blue) yields negative R^2, and that for ***Hs*** on MYI (red) is as low as 0.15. For example, for MYI, there exists large overestimation of ***Hs*** when ***Hs*** is small, and underestimation of ***Hs*** when it is large.

Table 1. R^2 for the least-squares fit (slope=1) for the retrieval.

	Retrieval with FB_{snow}, SMOS TB, and global ***s***	Retrieval with FB_{snow} and covariability (***alpha*** and ***beta***)
Hi in FYI	0.96	0.93
Hi for MYI	0.83	0.96
Hs for FYI	0.78	--
Hs for MYI	0.57	0.15

This result is in direct comparison to the retrieval carried out in the manuscript (Figure 7.c and d of the manuscript). Table 1 (above) shows the direct comparison of the R^2 for the least squares fit under the constraint of slope=1. It is shown that with the integration of SMOS TB, the quality of retrieval is improved mainly for ***Hs***. For ***Hi***, we consider the dominant factor is the value of freeboard, and the error in the retrieved ***Hs*** in Figure 3 (above) does not have big impact over the fitting for ***Hi***.

b) The retrieval as performed in this study is not very representative for an "actual" retrieval:

1. Only 50% of the available SMOS grid cells are used for the analysis, based on the criterion that "the error" (do you mean RMSE here?) between simulated and SMOS observed TB is < 1.5 K (as compared to 3.1 K for all SMOS cells). In a "real" retrieval situation where we do not have the information from OIB (i.e. hi, hs, and surface *temperature) to simulate TBs, how can we identify these cases where simulated and SMOS-observed TBs differ more?*

The authors would like to clarify that: (1) the retrieval with half of the points is not an inherent limitation of the proposed method (shown below in details), (2) the RMSE of TB for all the cells is 3.1 K, while the value of 1.41 K is the RMSE of cells with better altimetry (OIB) coverage, (3) the sufficient spatial sampling of altimetry (e.g., OIB) potentially results in better retrieval results, which can be detected through other methods or data (discussed below).

Since there is inherent discrepancy between the sea ice cover that generates the SMOS TB observations and that scanned by OIB (or any type of altimetry), we carry out the analysis of the TB error and its relationship to the coverage of OIB. To simplify the analysis, we use the OIB sample count (M) as the criterion for the spatial coverage of OIB. OIB campaigns include certain areas with extensive fly-over instead of a single line scan, and the effective sample count M is large for the corresponding cells. Figure 4 (below) shows the decrease of RMSE in TB (modeled v.s. SMOS) with the increase in M, analyzed over all available OIB data. For cells with larger M (i.e., over 95-percentile for M), the RMSE of TB drops from 3.1 K (for all cells) to 1.41 K, which is also mentioned in the manuscript (pg. 6, l. 5-6). Therefore, for the analysis in Section 4, we carry out retrieval for all cells with modeled TB within 1.5 K of the observed TB.

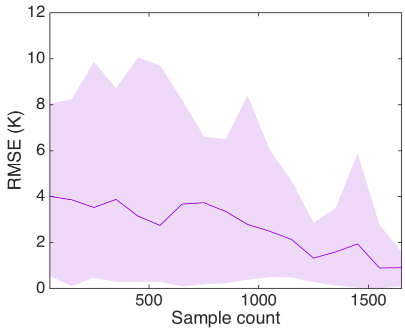
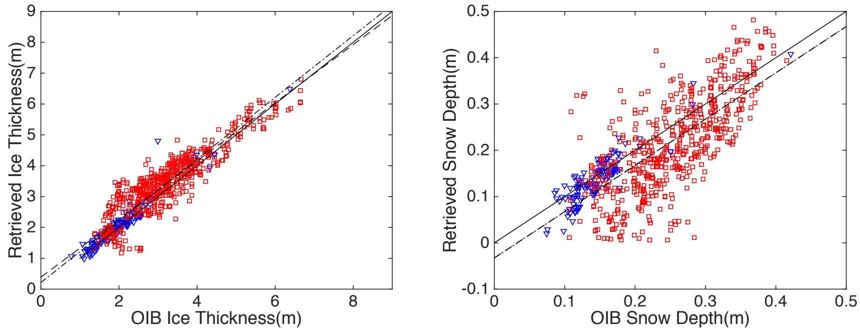


Figure 4. The relationship of RMSE of TB to OIB sample count.
For each sample count bin, the statistics of RMSE of TB are computed.
Shaded area covers the 5-th and 95-th percentile of the absolute TB error.

We have also carried out retrieval with ALL points (with SMOS TB and global value of ***s***). The verification for Hi and Hs are shown in Figure 5 (below). The quality of retrieval witnessed a slight drop when all cells are involved (R^2 of 0.89 to 0.86 for ***Hi***, and 0.64 to 0.56 for ***Hs***). Still, the both parameters can be retrieved.



(a) ***Hi***. (b) ***Hs***.

Figure 5. Verification with all points.

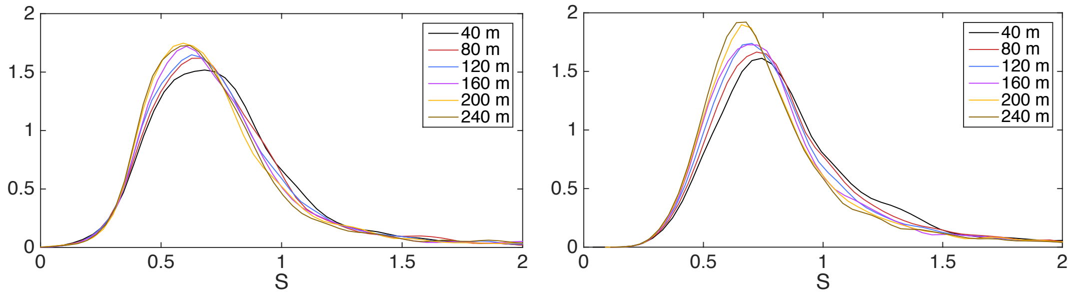
As demonstrated, the lack of sufficient spatial sampling of altimetry scanning plays an important role in the forward L-band model as well as the retrieval, and it is also very important to traditional satellite altimetry. With respect to the referee's question regarding the detection of potential TB errors, the authors consider utilizing methods and data that can support the detection of sub-cell variability of the sea ice cover (e.g., within 40km for SMOS). High-resolution remote sensing such as high-frequency radiometry (89 GHz channel from AMSR-E/2) or visible/infrared (onboard MODIS) serve as candidates. How they can be used for the study of representativeness of altimetry scans and the combined retrieval with L-band data, remains an important direction of future study.

2. In the retrieval here, TBs are simulated taking the surface temperature from OIB measurements. In a "real" retrieval situation, one would have to use other sources for surface temperature, which have different spatial and temporal resolutions; they could be (and most likely are) measured with many hours time shift! And as surface temperature has been shown to have a huge influence on L-band TB (see e.g. Maaß et al., 2013) and surface temperature can vary on short time scales, the results can be very different when using other temperature information. This is shortly mentioned at the very end of the manuscript, but it is not stressed how much this can influence the retrieval performance.

As pointed out by the referee, the surface temperature is generally not available for the retrieval. There are several third-party datasets that serve as candidates, including those based on remote sensing (such as MODIS and AMSR-E/2) and reanalysis data (e.g., ERA-Interim). Currently we are considering the retrieval on the daily basis. For example, the MODIS surface temperature field (of about 1 km of spatial resolution) on the same date of the SMOS TB is regridded to the retrieval scale of EASE grid (12.5 km), and further used for the retrieval. In a separate study, the MODIS data is adopted for the retrieval with radar altimetry (from CryoSat-2) and SMOS TB on a daily basis, and the retrieval generates good results (not shown here).
As noted by the referee, rapid change of surface temperature is present for the Arctic. However, we consider the rapid warming of the sea ice covered region is rare during winter (when altimetry is available), and mainly associated with sea ice leads, involving fast sea ice growth. As noted in Zhou et. al. (2017), the effects of open water and (refrozen) leads can be integrated in modeling of TB using third-party data (e.g., sea ice concentration and lead maps), by treating it as a small-scale variability of the sea ice cover.

3. The retrieval is done for freeboard measurements from the OIB campaign's laser altimeter. It is not discussed how the freeboard data from this altimeter compare with satellite based laser altimeters, which are given as the target for an "actual" retrieval.

The authors thank the referee for the comment on relationship with actual retrieval with satellite altimetry. The resolution of OIB is 40 m in the direction of the flight track, with the swath (across track) of about 100 m. As a comparison, for ICESat (or the to-be-launched ICESat-2) the resolution for each altimetry scan is about 70 m in diameter, with about 170 m between scans in the direction of flight track. Without considering the availability issues of laser scans (due to penetration problems caused by clouds), there exists resolution discrepancy between OIB and actual satellite measurements by ICESat.



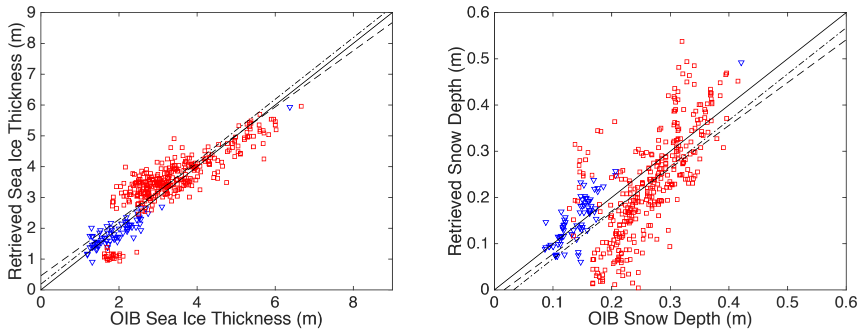
(a) FYI. (b) MYI.

Figure 6. Scaling of ***s*** derived from OIB data. PDFs of ***s*** are shown.

When working with satellite data, the characteristics of the covariability on the spatial scale of satellite altimetry has to be deduced. With respect to the covariability in the retrieval study, the issue is to compute the proper value of ***s***. We have carried out analysis by manually coarsening OIB's data to various spatial scales (by computing mean values for ***Hs*** and $\boldsymbol{FB_{snow}}$ with adjacent scans), and the results are shown in Figure 6 (above). There is a slight shift of ***s*** to smaller values for both FYI and MYI, when coarsening the OIB scans. Since ICESat(-2) features approximately 70 m spatial resolution, we intend to use the interpolated value of ***s*** between 40 m and 80 m for the potential retrieval with ICESat-2 and the concurrent satellite campaigns, such as WCOM.

c) With the current presentation of results I am not convinced that this is more than "playing around" with the relationship between ice thickness and snow depth (as observed during the OIB flights), which is taken advantage of. One would expect a relationship in that older ice is on average thicker and has had more time to accumulate snow on top, which leads to on average thicker snow on thicker ice (considering larger scales). And of course snow depth and freeboard are correlated as snow depth and ice thickness (together with snow and ice densities) determine the freeboard. Thus, I would not consider it as an "result" that you found a correlation between them. One interpretation (assuming that the densities of ice and water are relatively constant) is that the obtained R^2 value shows the fraction of the FB that is actually snow (R^2=0.53 would then mean that about 53% of the FB part consists of snow, while the rest is ice). This is also pointed out in the paper by Kwok et al. (2011), which is cited by the authors.

The authors would like to provide further proof that it is not statistics between ***Hi*** (or $\boldsymbol{FB_{snow}}$) and ***Hs*** that plays a central role in the retrieval. First, we ignore the covariability by carrying out retrieval with the mean value of $\boldsymbol{FB_{snow}}$. Specifically, for each previous retrieval problem (which involves multiple altimetry measurements and a single SMOS TB), we compute the mean of the *M* OIB samples of $\boldsymbol{FB_{snow}}$, denoted $\overline{\boldsymbol{FB_{snow}}}$. Then the retrieval is carried out between $\overline{\boldsymbol{FB_{snow}}}$ and TB, and the retrieved ***Hi*** and ***Hs*** are compared with the mean sea ice thickness and mean snow depth as measured by OIB (denoted $\overline{\boldsymbol{Hi}}$ and $\overline{\boldsymbol{Hs}}$). The results are shown in Figure 7 (below). This idealized retrieval scenario ignores the resolution difference between laser altimetry and L-band SMOS data. It is worth noting that no covariability is incorporated. But still good retrieval is achieved, with R^2 for ***Hi*** of 0.78 and that for ***Hs*** of 0.50.



(a) ***Hi***. (b) ***Hs***.

Figure 7. Retrieval with TB and mean $\boldsymbol{FB_{snow}}$. Dashed (dotted-dashed) line in each subfigure is the least-squares line (least-squares line under the slope=1 constraint).

Second, we further compare these results with the reported results in the manuscript, which is realistic retrieval involving multiple $\boldsymbol{FB_{snow}}$ measurements from laser altimetry. With the variability (i.e., *M* samples), the retrieval of ***Hi*** and ***Hs*** is improved, with R^2 rising to 0.89 and 0.64, respectively (see Figure 7.c and d in the manuscript). This indicates that the altimetry sampling provides necessary information for the retrieval accuracy. It also provides further clarification of the role of the statistics during retrieval: (1) the covariability between $\boldsymbol{FB_{snow}}$ and ***Hs*** DOES play an important role for the retrievability during the realistic retrieval, which involves relatively high-resolution altimetry scans, as studied in Section 3.2 of the manuscript, and (2) the statistical information DOES NOT play a central role in the general data fusion methodology of L-band TB and laser altimetry.

Comments B:

- Acknowledgements: OIB and SMOS data are not cited correctly here, see e.g. the NSIDC's and ICDC's (University of Hamburg) conditions/suggestions of how to cite the data usage

According to referee's suggestions, the acknowledgements is revised as below: "*This work is partially supported by National Key R&D Program of China under the grant number of 2017YFA0603902 and the General Program of National Science Foundation of China under the grant number of 41575076. The authors would like to thank the editors and referees for their invaluable efforts in improving the manuscript. SMOS data is provided from Integrated Climate Data Center (ICDC), icdc.cen.uni-hamburg.de, University of Hamburg, Germany, Digital media. http://icdc.cen.uni-hamburg.de/1/daten/cryosphere/l3b-smos-tb.html. [Data Accessed: 2017/10/25]. OIB and SSM/I sea ice concentrations data are provided by NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA. doi: http://dx.doi.org/10.5067/7XJ9HRV50O57. [Data Accessed: 2017/10/25]. Besides, the authors are grateful to Willmes, S. and Heinemann, G. for the provision of Arctic sea ice lead map.*".

- p. 3, L7f: "Several recent studies focus on the retrieval of snow depth over thick sea ice, based on L-band passive microwave remote sensing data from SMOS (Tian-Kunze et al., 2012)." -> I am not aware of "several" studies, they are also not given here, and the given reference (Tian-Kunze et al., 2012) is not about retrieving snow depth but (thin) ice thickness.

The authors recognize that the reference associated with snow depth retrieval is erroneous, and would like to make revision as below: "*The recent study in Maaß et al., (2013) has carried out the retrieval of snow depth over thick sea ice, by utilizing L-band passive microwave remote sensing data from SMOS*".

- p. 4, L17: As far as I know, "the OIB Level-4 product IDCSI4" that is claimed to have been used for 2012 to 2015 in this study is only available for the years 2009 to 2013... (While the IDCSI2 Quicklook data is indeed available for 2012 to 2015)

The authors recognize the reference here is incomprehensive and revised as below: "*Therein, the OIB Level-4 product IDCSI4 is adopted (Kurtz et al., 2013) for 2012-2013 and the remaining OIB data for 2014-2015 is from IDCSI2 Quicklook product, which is also available at NSIDC DAAC. Both of these two datasets are 40 m in resolution along the track's direction*".

- p. 4, L11-13: "...based on Burke et al. (1979). ... An adapted version of the model was adopted by Tian-Kunze et al. (2014) and ..." -> This is not correct. Tian-Kunze et al. (2014) use another (simpler) approach, which is originally based on a paper by Menashi et al. (1993). (also: "adapted version... was adopted" sounds strange)

The authors recognize that this is a mistake in the manuscript. The correct reference should only include Maaß et al., (2013a), which adopted the Burke's model. The whole sentence is then corrected as: "*In Maaß et al. (2013b), this model is applied to sea ice and further used for the retrieval of snow depth over thick sea ice*".

Comments C:

- p. 2, L17: "schematic view of remote sensing of sea ice" -> this seems a bit exaggerated to me, the figure mainly shows the definitions of snow and ice thickness and freeboard...

According to the suggestions of the referee, the sentence at p.2 L17 and the caption of Figure 1 are revised as follows: "*Figure 1 shows the various parameters related to satellite based laser altimetry and L-band passive radiometry for the sea ice cover*", and "*Sea ice parameters in the active and passive remote sensing of the sea ice cover, including sea ice thickness (hi), snow depth (hs) and snow freeboard (FBs)*".

- p. 2, L28-29: Not everyone knows what "its adapted version for ... FYI" is

The sentence is revised as follows: "... *climatological snow depth in Warren et al. (1999) for multi-year sea ice (MYI) and halved for first-year sea ice (FYI)*".

- p. 3, L9: Here, "near realtime" observations are compared with altimetry "which can only achieve basin coverage on the scale of about one month" -> It would be helpful to add that SMOS provides not only "near realtime" data but also "an almost daily coverage of the polar regions".

The sentence is revised as follows: "*SMOS provides full coverage of polar regions on a near real-time (daily) basis, observations with full coverage of polar regions. It has great advantage over satellite altimetry …*".

- p. 3, L13: "Despite the limited coverage..." -> strange wording/argumentation

The sentence is revised as follows: "*Although airborne remote sensing methods have limited spatial and temporal coverage, campaigns such as …*".

- p. 4, L7: "due to the limitation of satellite's orbital parameters, the inherent resolution is about 40 km" -> The resolution is determined by the antenna size, the frequency and the interferometry ("aperture synthesis") principle, not "orbital parameters".

The authors recognize the correction by the referee, and the sentence is revised as: "*However, due to the limitation of satellite's antenna size, the effective resolution of L-band radiometer onboard SMOS is about 40 km.*"

- p. 4, L25-26: "we consider OIB measurements in the adjacent 3x3 cells ... of equal contribution" -> Maybe mention at least that this is an approximation (the contributions are actually not equal, see the SMOS "antenna gain function")

According to the referee's suggestion to clarify the description, the authors have made two modifications. First, a sentence is added in the SMOS data description in Section 2.1: "*The gridded SMOS TB data field is generated from multiple snapshots within a day, with each snapshot involving multiple incident angles and spatially varying gain*". Second, the sentence mentioned by the referee is revised as "*However, due to the inherent resolution of SMOS is about 40 km and the daily gridded field is used in this study, we approximate the correspondence of OIB and SMOS TB by considering OIB measurements in the adjacent 3x3 cells (the red segment in Figure2) of equal contribution to the SMOS TB at the central cell (the one bounded by thick blue lines in Figure 2)*".

- p. 4, L30-31: "However, for certain segments of the OIB tracks, there exists extensive scanning which corresponds to a much larger value of M." -> I think it would be better and more precise to state the range of encountered M values (e.g. giving minimum, maximum and mean).

Figure 8 (below) shows the distribution of M for all available OIB data. The mean value of M is about 700. The referee is also kindly directed to Figure 4 (above) for the relationship between M and the RMSE of TB. Revision to the manuscript is also made to reflect the statistics of M.

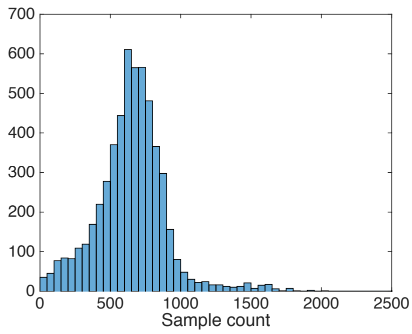


Figure 8. Distribution of OIB sample count (M).

- p. 5, L3-5: "The purpose of these treatments is to rule out the factors that may compromise the quality of the OIB samples and allow focus on the discussion of the retrieval algorithm." -> However, these conditions that were excluded here do not only influence the OIB data quality but will probably also make a potential retrieval with SMOS data more difficult and should be discussed somewhere.

The authors would like to point out that, indeed open water and (refrozen) leads can have profound effect on the overall TB, indicated by Kaleschke et al. (2010) and Tian-Kunze et al. (2014), and also studied in Zhou et al. (2017). Specifically, in Zhou et al. (2017), their effect is integrated into the radiation model by considering them (i.e., open water, leads) as a sub-scale mixture of the sea ice cover. Besides, a thermodynamic model for the diagnosis of ice thickness within the leads is adopted. For the basin-scale retrieval, these factors definitely should be considered, by using third-party data as indicators. But for the current study, we focus on the retrieval methods, therefore their effects are not considered. Outlook to how their effects should be accounted for is now included in Section 5 in the revised version of the manuscript.

- p. 5, L16: Reference to used radiation model should be given already here.
From the given reference it is not clear how the authors of that paper have "reformulated the model to include multiple layers for sea ice and snow" (Zhou et al, 2017). If several ice layers are used in the model, the higher order reflection terms should be considered. I did not find a statement on this...

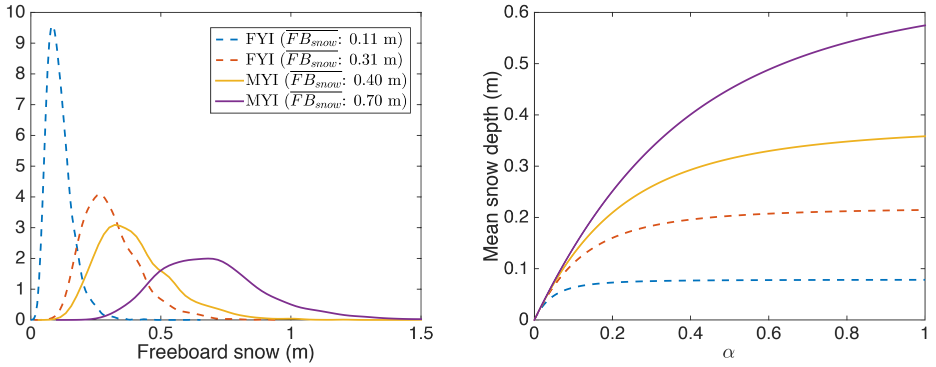
Following the suggestion of the referee, the authors add the reference to Zhou et al. (2017) to this sentence. The work for the study of radiation model is now currently available online (see Zhou et al. (2017)). The non-coherent model from Burke et al. (1979) and its adaptation to sea ice in Maaß et al. (2013a) is the basis for the study in Zhou et al. (2017). Since much higher overestimation of TB was witnessed for MYI, in Zhou et al. (2017) the drainage of salinity in the top part of the MYI is accounted for, by using a multi-layer formulation of the model. We provide a model description document as a supplement to the revised manuscript.
The original model in Burke et al. (1979) ignores the high-order terms for reflection (beyond 2^{nd} order). As reported in Maaß et al. (2013b), there exists different behavior of the multi-layer of Burke's model and the multi-layer coherent Ulaby model. However, two aspects that are relevant to the current study: (1) as compared with the drainage of top-layer salinity, the multi-layer formulation induces a smaller change as compared with the single-layer formulation (adopted by Maaß et al. (2013a)), and for MYI, the integration of salinity profile resulted in a much better fit of modeled TB to SMOS observation, as shown in Zhou et al. (2017); and (2) as shown in Figure 3.c of the manuscript, the potential of retrieval lies in that the retrieval is carried out over constant freeboard lines, and even if TB saturates (beyond 3 m for ***Hi***) there still exists good sensitivity for retrieval.

- p. 7, Eq. 3: Maybe better to use "arctan(...)" because tan^-1(...) could also be interpreted as 1/tan(...)

Corrected according to the referee's suggestion. All other cases including "*tan^-1*" in the manuscript are corrected to "*arctan*" as well, in order to avoid misinterpretation and ensure consistency.

- p. 8, L5 - p. 9, L19: First you write about "scanning of alpha" without explaining it at all, then you present the results shown in Fig. 5. Then you explain the "scanning of alpha" procedure and finally refer to Tab. 2. This is very confusing. The scanning of alpha should be explained first. It would also be very interesting to see what values alpha takes in this procedure. Are they similar for the individual retrievals? Are they spread over a large range? Do the results shown in Fig. 5 and Tab. 2 belong to the same analysis?

According to the suggestions of the referee, we have made revisions to the manuscript: (1) first we include the description of ***alpha*** and the scanning of it in Section 3.2 (on p. 9), (2) in Section 4, the scanning of ***alpha*** is introduced in the retrieval process in more clear way. Figure 1 and 2 of this response document also provide the statistics of ***alpha***. As shown, with OIB data, there exists a large range of possible values for ***alpha***, which is also indicated by the retrieval process (which generates the value of ***alpha*** as well). In Figure 9 (below), for the typical distribution of FB_{snow} for both FYI and MYI (subfigure a), the scanning of ***alpha*** and the resulting mean ***Hs*** is shown (subfigure b). Since the value of ***s*** is smaller than 1, there exists no inundation, which is consistent to the rarity of such events in the Arctic. Besides, indeed, Figure 5 and Table 2 of the manuscript belong to the same analysis.



(a) $\boldsymbol{FB_{snow}}$ distribution. (b) Mean $\boldsymbol{Hs}$.

Figure 9. Scanning of parameter α and the corresponding mean snow depth ($\boldsymbol{Hs}$).

- p. 10, L3-6: "There is minor increase in quality (0.91 versus 0.89 and 0.65 versus 0.637) and a relatively large gap to the "ideal" case. This indicates that the uncertainty (or error) in TB and radiation models plays an important role in affecting the quality of the retrieval.
The uncertainty of TB may arise from that of the radiation model, as well as the mismatch between the altimetry and passive microwave remote sensing" I think the very similar results for simulated and SMOS TBs cannot lead to these conclusions about the radiation model or the TB measurements! Even if (theoretically) the radiation model gave completely unrealistic TB values: If you do the hs/hi-retrieval by comparing this radiation model's TBs (for different hs,hi values) with this same radiation model's "true" TB (for the "true" hs,hi values), the difference in retrieved and "true" hs,hi will originate from other assumptions used in the retrieval (here: assumption of Eq. 3, choice of s-value, choice of thresholds, ...) or from the ill-posedness of the problem (TB ambiguities in the model, which can also exist in reality) but not the quality of the model to represent the "real world"/SMOS (because you are comparing it with its own output! You are within its "ideal model world"). In contrast, an existing difference between using SMOS and simulated TBs may contain information on the radiation model's performance to simulate SMOS TBs (and also on the effect of the spatial mismatch between altimeter and satellite measurements).
As far as I can see, the difference between using global and local s values tells you something about how good the global s value approach is. Here (with the results for simulated and SMOS TBs being very similar), THIS (=using different s values) is where the "relatively large gap to the 'ideal' case" seems to come from!

After careful re-check of the manuscript and the results, the authors have discovered a mistake in the manuscript, and would like to sincerely apologize for the error and potentially misleading the referee. The retrieval result with R^2 of 0.91 for $\boldsymbol{Hi}$ and 0.65 for $\boldsymbol{Hs}$ is attained with the local value of $\boldsymbol{s}$ and SMOS TB, and NOT with the global value of $\boldsymbol{s}$ and the modeled TB. In the manuscript (pg. 10, l. 1-3), the correct sentence should be: "*Furthermore, if the retrieval is based on: (1) observed TB from SMOS, and (2) the locally fitted value of* ***s****, the* R^2 *values for the fitting are 0.91 and 0.65 for sea ice thickness and snow depth respectively, with virtually no change in the fitting lines (not shown)*". For the sake of completeness, we have also carried out retrieval with the other combination of modeled TB and global value of $\boldsymbol{s}$, and the R^2 values are 0.96 and 0.84 for $\boldsymbol{Hi}$ and $\boldsymbol{Hs}$, respectively. These values are very close to the "ideal" retrieval case (which involves modeled TB and the local value of $\boldsymbol{s}$). All the results for fitting (for all 4 combinations) are shown in Figure 10 (below). These results indicate that the difference between modeled and observed TB is indeed a major source of the retrieval error, especially for $\boldsymbol{Hs}$, which is a "correctly" stated sentence in the manuscript (i.e., the sentence the referee mentioned).

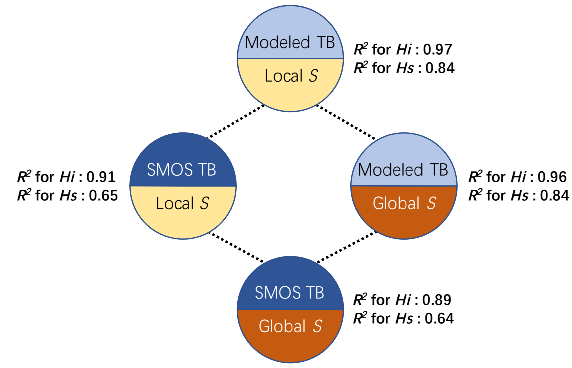
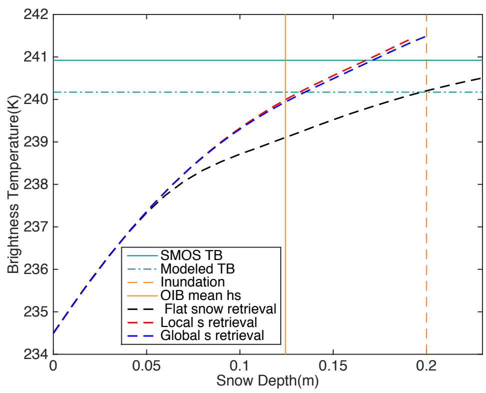


Figure 10. Comparison of different configurations in retrieval. Each circle represents the retrieval with the specific TB and the value of ***s***. The bottom circle represents the realistic retrieval scenario, and the top one represents the "ideal" scenario. The fitting to observations (R^2) for both ***Hi*** and ***Hs*** are shown accordingly.

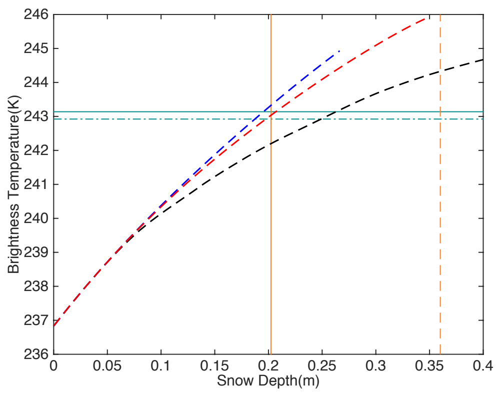
The authors would like to emphasize again that the even if with the mean freeboard and TB, both ***Hi*** and ***Hs*** can be retrieved (shown in Figure 7 above). However, when the information of altimetry samples is incorporated, better retrieval can be achieved for both ***Hi*** and ***Hs***. For the retrieval which accounts for the resolution difference between L-band TB and laser altimetry, covariability plays an important role in the retrievability.

- Fig. 5 would benefit from a legend and/or annotations of some of the lines, it is hard to remember from the figure caption what each of the 7 lines represents...

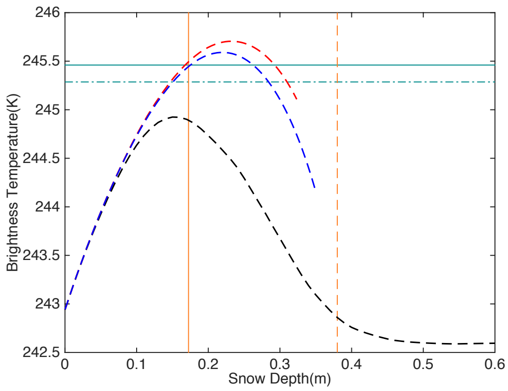
Legends are added to each of the subfigures of Figure 5 for better readability. The updated figure is available in the revised manuscript. They are also reproduced below in Figure 11.



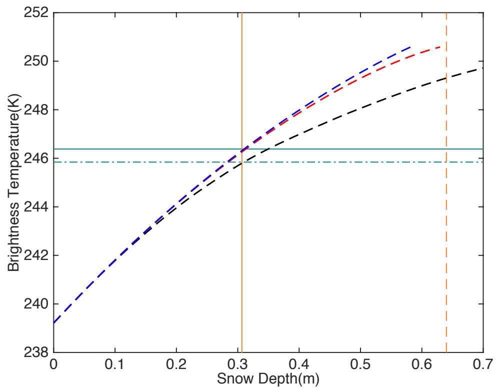
(a) Scenario I



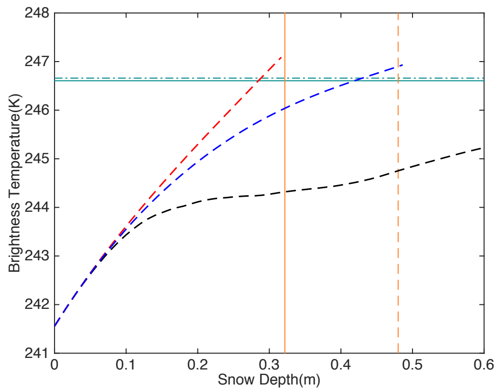
(b) Scenario II



(c) Scenario III



(d) Scenario IV



(e) Scenario V

Figure 11. Updated Figure 5 of the manuscript.

Comments D (not a complete list!):

- Usage of "etc" is not very precise, for example on p. 2, L5 & p. 2, L20 & p. 4, L11 & p. 5, L34

The authors have made revisions to the manuscript for corrections and more precise use of “etc”, including those pointed out by the referee. Unnecessary uses of “etc” are deleted.

- p. 2, L5-6: "there is rapid" -> "there has been rapid"

Corrected.

- p. 2, L24: "hence limited spatial coverage" -> not a complete sentence

Corrected as: “*..., so they are limited in terms of spatial coverage*”.

- p. 3, L6: "researches...obtain snow depth" -> strange wording

This sentence is revised by segmenting into three parts: “*The retrieval of snow depth with passive microwave satellite remote sensing has been carried out in various studies. In Comiso et al., (2003), multi-band data from AMSR-E are utilized, but only for FYI. Maaß et al., (2013b) explored the retrieval of snow depth over thick sea ice with L-band data from SMOS*”.

- p. 3, L10: "requires the prerequisite" -> either "requires" or "prerequisite"

Corrected as follows: "*However, the sea ice thickness is required for the retrieval*".

- p. 3, L25: "achieves successfully retrieval" is not a correct expression

This sentence is revised as follows: "*..., we demonstrate that the proposed algorithm can simultaneously retrieve both sea ice thickness and snow depth, ...*".

- p. 3, L26: "correspond" -> "corresponds"

Corrected.

- p. 4, L20: "Temporally, the date of each OIB campaign is located, and the SMOS TB data from the specific date is attained for the combined retrieval." -> An example for a case where the readability could be improved. This sounds like a complicated way of saying something like: "OIB and SMOS data from the same day are taken."

Corrected as indicated by the referee as follows: "*OIB and SMOS data from the same day are taken.*"

- p. 4, L23-24: "due to the inherent resolution of SMOS data is about 40km, therefore..." is not a correct expression

The sentence is corrected as follows: "*However, since the inherent resolution of SMOS data is about 40 km, even if the SMOS data product is provided on the 12.5 km resolution (small blue cells in Figure 2), we consider OIB measurements in the ...*"

- p. 4, L26: "the 9 cells covers" -> "the 9 cells cover"

Corrected.

- p. 4, L28: "the total area the contributes" -> "the total area that contributes"

Corrected.

- p. 6, L28 - p. 7, L7: This part is hard to understand...

The authors would like to explain that: the purpose of this paragraph is to introduce the protocol in the covariability analysis. For clarity, the whole paragraph is revised as follows:

"*For the covariability between* $\boldsymbol{FB_{snow}}$ *and* $\boldsymbol{Hs}$*, we choose the native resolution of the OIB product (40 m) as the spatial scale for analysis. Each TB corresponds to multiple (M) OIB samples, with each sample containing the measurement for both* $\boldsymbol{FB_{snow}}$ *and* $\boldsymbol{Hs}$*. We divide these samples into* $\boldsymbol{FB_{snow}}$ *bins, with each bin covering 5 cm. In total there are 30 bins, covering the range of 0 to 1.5 m. For samples in each bin, we compute the percentiles and the mean value of* $\boldsymbol{Hs}$*. Figure 4.a shows the mean* $\boldsymbol{Hs}$ *and the +/-1 standard deviation range and their relationship with* $\boldsymbol{FB_{snow}}$*, for 4 representative TB points. Furthermore, we carry out least squares fitting (weighted according to sample count in each bin) between mean* $\boldsymbol{Hs}$ *and* $\boldsymbol{FB_{snow}}$*. Among all available TB and OIB data, there exist statistically significant positive correlation between* $\boldsymbol{Hs}$ *and* $\boldsymbol{FB_{snow}}$ *for over 90%*

of all points. The values of R^2 *are in the range of 0.06 and 0.89 (95% percentile), with the mean value of* R^2 *of 0.53. This indicates that there exists consistent covariability between snow depth and snow freeboard across Arctic sea ice cover*."

- p. 8, L1-2: "For hs > 0, there will be inundation due to: FB< hs." -> Do you mean: "For hs> 0, there will be inundation for values hs>FB."?

The referee's guess is right. For the sake of clarity, this sentence is revised as: "*For any sample of* $\boldsymbol{FB_{snow}}$*, if the value of freeboard is smaller than the current value of* ***hs****, in order to avoid inundation, the snow depth for this sample is assumed to be the same as* $\boldsymbol{FB_{snow}}$".

The authors would like to express sincere thanks to the referee for the insightful comments and invaluable suggestions to the manuscript. Revisions to the manuscript are made according to the suggestions. We hope that through the reply and the revised manuscript, the idea and results are now better conveyed to the referee. We would also like to answer any further questions and comments from the referee.

References:

Kaleschke L, Maaß N, Haas C, et al. A sea-ice thickness retrieval model for 1.4 GHz radiometry and application to airborne measurements over low salinity sea-ice[J]. The Cryosphere, 2010, 4(4): 583-592.

Maaß, N., Kaleschke, L., Tian-Kunze, X., and Drusch, M., 2013a: Snow thickness retrieval over thick Arctic sea ice using SMOS satellite data, The Cryosphere, 7, 1971–1989, doi:10.5194/tc-7-1971-2013, www.the-cryosphere.net/7/1971/2013/.

Maaß N, Kaleschke L, Stammer D. Remote sensing of sea ice thickness using SMOS data[D]. University of Hamburg Hamburg, 2013b.

Tian-Kunze, X., Kaleschke, L., Maaß, N., Serra, M. M. N., Drusch, M., and Krumpen, T.: SMOS-derived thin sea ice thickness: algorithm baseline, product specifications and initial verification, The Cryosphere, 8, 997–1018, doi:10.5194/tc-8-997-2014, www.the-cryosphere.net/8/997/2014/, 2014.

Zhou L, Xu S, Liu J, et al. Improving L-band radiation model and representation of small-scale variability to simulate brightness temperature of sea ice[J]. International Journal of Remote Sensing, 2017, 38(23): 7070-7084.