

Response to the reviewers

‘On retrieving sea ice freeboard from ICESat laser altimeter’

Kirill Khvorostovsky, Pierre Rampal

Nansen Environmental and Remote Sensing Center, 5006 Bergen, Norway

5 *Correspondence to:* K. Khvorostovsky (kirill.khvorostovsky@nersc.no)

Dear Referees,

Thank you very much for very careful review of the manuscript and many useful comments helping us to improve our paper.

Please find below the answers to your comments and *the changes we have made in the manuscript*:

1 Response to Referee #1 (Stefan Kern)

10 General comments

1

I have the feeling that the authors could work on the motivation and the principal aim of the paper. What are the new findings of this paper? How relevant are these? How globally applicable are these? Since ICESat-1 is not in orbit anymore one could ask whether it is worth to apply the new approach presented and what a potential user would gain from that. To me it seems as if the main improvement is 1-2 cm smaller bias for basically first-year ice (FYI). This might be not enough to trigger a potential user to switch to the new, optimal approach. One could, however, use this paper to give another, additional evidence for how difficult it is to derive sea ice thickness from satellite laser altimetry.

15 In the revised manuscript we have tried to highlight main findings of the paper and to emphasise that the aim of the paper is to analyse the different sources of the uncertainties in freeboard retrieval using ICESat data. In particular, we rewrote most part of the abstract, changed and extended discussion of the results following, in particular, reviewer’s comments, and added a paragraph in the conclusions.

2

25 A relatively large number of the results roots on the QuikSCAT data based discrimination between FYI and multiyear ice (MYI). The authors used a 50% MYI (or FYI) isoline to separate between both ice types. I have the feeling that the authors could enhance their results by reducing the number of mixed pixels and instead of using 50% defining FYI and MYI areas as follows: FYI: all grid cells with MYI fraction below 10%; MYI: all grid cells with MYI fraction above 90%. I could imagine that the difference in the freeboard and in the effect the authors are

describing becomes even more clear in that case. And in addition the authors would concentrate their ice-type based analysis on grid cells where the discrimination into these surface types is more reliable and the relative error contribution is smaller than when using a 50% threshold. Yes, by this the sum of the grid cells analysed (FYI plus MYI) would not be the total anymore, but I don't think that this would have a significant influence on results and interpretation.

We have checked how the mean freeboards will be changed for FYI and MYI in case of using 10% and 90% thresholds respectively instead of one threshold of 50%. However using of these new thresholds does not noticeably change the difference between the results before and after applying the improved algorithm. In order to emphasize the importance of proposed improvements in the TP method we additionally note their effect to the FM08 period when significant freeboard differences are observed.

3

The authors test two different, constant gain settings. I am wondering whether they are aware of the paper by Yi et al., 2011, Annals of Glaciology, where it is stated that due to the instability in the power emitted by the GLAS a constant gain setting might not be appropriate. Yi et al. (2011) suggest to use a variable gain setting - which was also used by Kern and Spreen, 2015, for instance. Donghui Yi was one of the main principal scientists working with ICESat data in the background and I doubt that there are many more who have more experience with the nitty-gritty technical details of this sensor than Donghui.

Indeed, Yi et al., (2011) suggest using varying gain settings for data filtering. In the revised manuscript we have mentioned it in the discussion of freeboard sensitivity to data filtering in section 3.1. However, both, the GSFC and JPL, products were derived by applying a constant threshold as stated in the descriptions of the freeboard retrieval algorithms (i.e. Yi and Zwally, 2009 and Kwok et al., 2007). Since the goal of the paper is to reproduce and compare freeboard retrieval methods used for generation of these two products, we applied constant settings. In this study we focus on the assessment of the effect of different approaches for determination of the sea level references and on the reasons of the differences between the GSFC and JPL products.

4

All maps seem to be blurry compared to the histograms. I am wondering whether the authors could increase the dpi for these maps. In addition, the color legends in all maps would benefit from a title and a unit.

The maps were originally produced as png files and their quality remains fine when inserted in the Word document. But, on the stage of creating the pdf document the maps becomes blurred. We suppose that this problem can be solved by providing the original png files or Word document.

We have added the units to the color bars of the maps.

Specific comments

Page 1, Line 13: "by up to 15 cm". In the next line the authors give a mean and an "up to" value. Therefore I would find it fair to see a mean value here as well.

5 We have rewrote most part of the abstract to present our findings in the paper more clear. In the new version of the abstract we do not give a mean values for freeboard differences separately for the FYI and MYI.

Page 1, Line 15: Perhaps add "average" before "difference"?

Agree, it has been added.

10

Page 1, Line 19: I would not write "is very large and". Here it comes to the main message the authors wish to give to the readership, i.e. whether the authors wish to give a list of potential uncertainties and biases or whether the authors wish to promote their improvement.

15 Agree, we have removed these words. In the revised manuscript we also rewrote most part of the abstract to present our results more clear.

Page 1, Line 23: One could argue that this statement is not so new but I am wondering whether authors could take a look the parameters used for the freeboard-to-thickness conversion. Doesn't the JPL approach use a specifically designed ECMWF data based snow depth data set?

20 We do not aware of any published studies that analyze the reasons of the difference between JPL and GSFC products. The question whether it is caused by the difference in the methods of freeboard retrieval or in the parameters used to convert freeboard to thickness or both.

Indeed, the GSFC and JPL products were generated using different approaches to determine snow depth. For the JPL product the ECMWF data were used, while for the GSFC product they used Warren climatology modified by applying empirically defined factors, which depend on ICESat period and freeboard values. However, we believe that the effect of the methods for freeboard to thickness conversion deserve separate analysis, while this study is focused on the discrepancies of freeboard retrieval algorithms.

Page 2, Line 2: I am wondering whether the authors would like to weaken that "ten times" by an "approximately". Currently this contradicts with the notion given on Page 12, lines 30- 32.

30 Agree, it has been corrected.

In addition, the words 'sea ice freeboard' used in the section 3.5 (Page 12, lines 32 of the initial version of the manuscript) have been changed to 'total freeboard'.

Page 2, Line 1-5: Even though the authors refer to the Zygmuntowska et al (2014) paper later I suggest that it could be mentioned here as well.

Agree, the reference has been added in this paragraph.

- 5 **Page 2, Line 8: I am aware of this 2 cm, but wasn't this obtained for a very ideal case of sea ice (smooth, extremely level young ice in the Ross Ice Shelf polynya)? I am wondering whether the authors also would like to mention the single-shot accuracy of 13.8 cm.**

Agree, this accuracy has been mentioned with corresponding reference.

- 10 **Page 2, Line 27: "their mean represents" I don't think this is entirely correct. If within a ground track segment of 550 km the LLE provides 3 elevations then these are used to approximate the SSH ... maybe by simply computing a mean in the Yi and Zwally (2009) version of the LLE but not in the version proposed by Spreen et al. (2006) and Kern and Spreen, 2015, where the SSH at that location is approximated by fitting a polygon through the elevations identified as minima.**

- 15 Indeed, since in our study we reproduced the LLE method used by Yi and Zwally (2009) for generation of the GSFC product, we focused on the description of their method, which uses mean of the lowest elevation. However mentioning of the fitting method with corresponding references has been added in this paragraph in the revised manuscript.

- Page 2, Lines 28/29: This statement is true for the missing leads but if the LLE is set to, e.g. 2% and there are fewer elevations than a certain predefined number (I guess 3 in Kern and Spreen, 2015; how is this solved in Yi and Zwally?), then for that specific laser shot no SSH is approximated and it will be instead estimated via interpolation from the neighboring estimates. The fact that a moving ground track segment is used here ensures that inconsistencies / jumps are smoothed along track.**

- 25 The LLE method, as applied by Yi and Zwally (2009) for the Arctic sea ice, implies using 1% of lowest elevation within running 100-km segment. Thus, they estimate local sea level for all laser shots (remained after data filtering). Therefore, in case if there are no leads within the along track segment the selected lowest elevations would measure the height of thin ice. This has been stated, in particular, by Zwally et al., (2008), where LLE method was proposed, and in the description of the GSFC product (Yi and Zwally, 2009).

- 30 **Page 2, Line 34: "particular to the chosen value for ice density": Is this true? I thought that Kwok (et al.), I guess in their 2007 paper, carried out a sensitivity / uncertainty analysis and figured out that in fact the snow depth is the second most important error contribution, after the uncertainties in freeboard.**

That is correct. We have modified this paragraph as follows: *The Arctic sea ice thickness from two available products derived from ICESat data by Jet Propulsion Laboratory (JPL) using the TP method (<http://rkwok.jpl.nasa.gov/icesat>) and by*

5 Goddard Space Flight Center (GSFC) using the LLE method (Yi and Zwally, 2009) were found to be different by 0.42 m (Lindsay and Schweiger, 2015). This difference can be caused by the different techniques for determining the local sea level in the freeboard retrieval algorithm, or by the different methods in estimating snow depth that is used for calculating ice thickness, i.e. by the uncertainty of the freeboard-to-thickness conversion (Kwok and Cunningham, 2008; Zygmuntovska et al, 2014).

Page 3, Line 10: "is also quantified": I am wondering whether the authors could state which version of the JPL products they reproduce: the simple first version or the optimized one. It is not clear from this paragraph.

10 We reproduce both versions of JPL products in order to assess, from the one side, the effect of using different methods on determination of sea level references, and, from the other side, differences between the freeboards used in the JPL and GSFC products. This sentence has been changed for clarification: *These two corrections were taken into account in the JPL product. Therefore we quantified their effect on freeboard estimates, and hence on the difference between the corresponding sea ice thickness products.*

15 **Page 3, Line 13: I suggest to add "elevation" behind "level 2"**

We are not sure that it would be correct to add "elevation" behind "level 2". Indeed, elevation is the primary product of altimeter data, but they also contain different signal parameters used for data filtering and correction.

Page 3, Lines 13-15: Please explain the meaning of "FM" and "MA" as well.

20 We added a sentence with description of these abbreviations denoting ICESat periods: *The abbreviations ON, FM and MA mean October-November, February-March and March-April respectively, followed by the year (i.e. 2003 to 2008).*

Page 3, Line 16: Is the Yi and Zwally data set also a level 2 data set? The authors could add this information.

25 Level 2 altimeter data contain only elevation measurements with corresponding parameters, while the Yi and Zwally data is the final product with the freeboard and thickness estimates. We added word "*product*" in order to emphasize the level of the Yi and Zwally data.

30 **Page 3, Lines 18-21: Please note on which grid projection these two data sets are given. While the NSIDC data set is publicly available - and may even have a DOI? - the one based on QuikSCAT seems to be an in-house (NERSC) product. Is this correct? Would it be possible to give more information about this data set? At the end this is relatively crucial for your work because a lot of the results you give are based on the discrimination of FYI and MYL.**
For more detailed description of these data this sentence has been changed to: *In addition, we use the Arctic-wide multiyear ice fraction dataset used in Zygmuntowska et al. (2014) that was produced by reprocessing the QuikSCAT satellite scatterometer data. Zygmuntowska et al. (2014) used daily averaged gridded (22.5 km) data of radar backscatter processed*

by Brigham Young University (<ftp://ftp.scp.byu.edu/data/qscat/SigBrw>), and converted them into multiyear ice fraction following the method described in Kwok (2004).

Page 3, Lines 24/25: This sentence is not clear to me.

- 5 We have reformulated this sentence to make it clearer: *When reproducing the TP and LLE methods, we used the same filtering criteria that are applied in Kwok et al. (2007), in order to compare the algorithms avoiding biases associated with different filtering.*

Page 4, Line 2: See general comment 3

- 10 As we noted in the reply to the general comment 3 we applied a constant setting following the description of the freeboard retrieval algorithms used for generation of the products.

Page 4, Line 12: I suggest to write: "... level in leads we evaluate the effect of using a different geoid on the results in section 3.2" instead of using passive voice.

- 15 Agree, the sentence has been changed.

Page 4, Line 17/18: I suggest to write: "We evaluate and discuss the effect of ..." instead of using passive voice.

Agree, the sentence has been changed.

- 20 **Page 4, Lines 16 & 25: I am wondering whether the authors would consider to add a table in which they list the different approaches, different scales and different settings used. I can imagine that this would increase the readability of their paper.**

Agree, the table with different scales and settings used for the JPL and GSFC products has been added as Table 1. Please, note that following the comment from the reviewer #2 we have also added a flow chart (as the Figure 1) with the methods and settings implemented in our study.

- 25 Please, note that due to adding of new Table 1 and Figure 1 in the revised manuscript the numbering of figures and tables is shifted.

Page 4, Line 24: I suggest to write: "... availability. The distance between ICESat along track samples is 172 m. If we

- 30 **..."**

Agree, the sentences have been changed.

Page 4, Line 25: It is just a minor detail but I suggest to write "about 580" instead of "about 600".

Agree, the number has been changed.

Page 4, Line 29: I suggest to give more information about the type of satellite images used by Kwok et al.

We have changed the beginning of the sentence to specify the type of images: *From the analysis of SAR images from RADARSAT satellite, Kwok et al. (2007) found...*

5

Page 5, Line 5: What is \bar{R} overbar?

In the end of the sentence a definition of \bar{R} has been added: ... *with \bar{R} being the mean reflectivity of all the measurements within 25 km around that given sample.*

10 **Page 5, Lines 9-11: I have difficulties to understand this sentence. What is meant with "their agreement"?**

The phrase "their agreement" means agreements of "these two sets of tie-points" mentioned in the beginning of the sentence. We have tried to reformulate this sentence to: *They found good agreement between two sets of tie-points as well as agreement of these sets with high-quality tie-points determined from collocation with satellite images.*

15 **Page 5, Lines 11-12: I have difficulties to understand this sentence as well. Why is the number of tie points not sufficient for basin wide studies?**

The tie-points for the first set are identified as samples with $\Delta R > 0.3$ and located below the regression line. This additional requirements of the dip in reflectivity limits the number of identified tie-points and their sampling density. We have slightly changed the sentence to note it: *Since a sampling density of tie-points in the former set is not sufficient for basin-wide studies*

20 ...

Page 5, Lines 14-15: I suggest to write: " ... the surface roughness of a given sample ..."

Agree, the sentence has been changed.

25 **Page 5, Line 15: What is meant by "to be for this sample"?**

We changed this sentence to make it clearer: *The higher σ_{25} , which is characteristic of the surface roughness of a given sample, the lower is the h_r is required to qualify this sample as a tie-point.*

Page 5, Lines 18-19: I suggest to write: "We discuss the influence of the regression model in Section 3.3"

30 Agree, the sentence has been changed.

Page 6, Lines 1-3: I am aware of this correction. In the frame-work of this paper and these studies, I suggest the authors could spend 1-2 sentences more on this issue. The TP method is the one you are finally after here. Now, if the TP method has - as one of the caveats - that it identifies snow covered, refrozen leads as potential tiepoints, then one

could doubt its reliability. Thin ice, to carry snow on it, needs to have a certain thickness. Could it be that this correction is also done to take the effect of frostflowers growing on thin ice into account. These increase the reflectivity as well. In this context the authors could include also a sentence or two into the discussion because this is an obvious limitation of the TP - or in general of methods using laser altimetry where the surface reflectivity of the sea ice plays a major role in the SSH approximation.

Two sentences about limitation of the function used to determine the correction have been added: *It can be noted that increase of reflectivity over young ice may also reflect the effect of frost flowers growth. Therefore, the function of reflectivity that account for accumulated snow and used to determine the correction may be different in the presence of frost flowers.*

10

Page 6, Lines 7-8: I suggest to formulate this more clearly in the last sentence of this paragraph. The correction does not increase directly with freeboard height because the latter is not part of the correction factor. Therefore it might be better to write about a relative increase. In contrast, the reflectivity R is part of the correction factor and directly influences it.

15 The correction does increase directly with freeboard. As we describe in the text, Kwok et al. (2009) calculated the correction as a product of freeboard and a factor $1.1 + 0.1 \left(\frac{R_{snow} - R}{R_{ice}} \right)$. Thus, the first term of this factor (i.e. 1.1) implies that this correction increases all freeboards by at least 10% of freeboard value.

Page 6, Line 29: Comments to Table 1: Does this table include values from the entire maps shown in the various figures or did the authors focussed on a certain region? I am asking because the work of Kwok et al. usually focusses on the Arctic Ocean, omitting seas such as the Hudson Bay, Greenland Sea, etc. I am wondering whether it would make sense to do the same in the present study - and if not what could be a good reason to not limit the area. If the authors keep the region like they did, then I suggest that they mention this in the discussion of the results when they are referring to differences between the different methods in terms of the retrieved sea ice thickness.

25 Indeed, the JPL product does not contains estimates over the seas surrounding the Arctic Ocean (in addition to Hudson Bay and Greenland Sea these are Barents Sea, Kara Sea and Baffin Bay). We believe that it is reasonable not to exclude these areas from the consideration in the sections where we analyze spatial distribution of the difference between the freeboards. Accordingly, we think that it is better to present the average values corresponding to the entire maps in the tables with the differences. However, in the new table proposed below by the reviewer (Table 4 in the revised manuscript) we present the average freeboards that correspond to the areas included in the JPL product. This makes sense because in the section 3.5 we discuss the differences in terms of their effect on the retrieved sea ice thickness and the difference of 0.42 m between the JPL and GSFC products obtained by Lindsay and Schweiger (2015), which corresponds to the Arctic Basin. For consistency, we use corresponding values also in the Figure 9 updated accordingly (Figure 10 in the revised manuscript).

30

Page 6, Line 31: "differences between data releases" I suggest the authors try to be less global here and to give an approximate estimate of a) what causes these differences and b) how large these are. This might also help to explain why the standard deviations of line 1 in Table 1 are relatively large.

5 We suppose that different data filtering is the main factor that causes the freeboard differences in the line 1 in Table 1 (Table 2 in the revised manuscript), and discuss it in more details in the section 3.1. Furthermore, the data from previous releases are not available anymore and it is not possible to provide detailed analysis of the differences between data releases. However, from the description of previous data releases one can presume that one of the most valuable improvements that has been made regards to saturation correction. Therefore, we mentioned this factor in the changed sentence and emphasized
10 that the differences between data releases is secondary factor: *The remaining discrepancies can be attributed to different data filtering, and possibly to the differences existing between data releases (e.g. improvements in saturation correction).*

Page 7, Line 1: "up to 10 cm" Just a comment: This is indeed a lot for sea ice thinner than 15 cm. This is a bias of up to 2/3 of the sea ice thickness and in terms of sea ice thickness derived from the biased freeboard this would amount to an up to 300% relative error. That is something.

Yes, indeed. This illustrate how the choice of filter settings contribute to uncertainty of the results.

Page 7, Lines 2-4: See my general comment 3.

As we noted, in this study we reproduce and compare freeboard retrieval methods used for generation of the GSFC and JPL
20 products, which are derived by applying a constant threshold for the receiver gain parameter. In order to point out to an alternative approach for applying gain setting we have added the following sentences in this paragraph: *It should be noted that Yi et al. (2011) used different thresholds for gain in order to account for the reduce of gain with age of the ICESat lasers due to decrease of the transmitted power. However, in this study, we compare freeboard estimates with the GSFC product derived by Yi and Zwally (2009) using constant setting for the gain threshold.*

25

Page 7, Line 13: "This explains ..." It seems a bit strange that the usage of a constant (possibly wrong) gain is selectively causing biases for specific ICESat periods. Again see general comment 3.

Since Yi and Zwally (2009) apply a constant gain threshold for the GSFC product, it cannot be the reason of the observed difference. At the same time, pulse-broadening parameter also may depend on ICESat's laser age. We have added the
30 following sentences for clarification: *The largest biases and their variability observed in the periods ON03 and ON04 correspond to the first operation periods of laser 2 (campaign 2a) and laser 3 (campaign 3a) respectively. Therefore, one may presume that the pulse-broadening parameter applied by Yi and Zwally (2009) is also affected by the instability in the power transmitted by the ICESat's lasers.*

Page 7, Line 24: The authors could add that the main purpose of this along-track averaging is to get rid of the large-scale fluctuations in the elevations caused by the geoid used.

The purpose of the along-track averaging is described in the first sentence of section 2.3, where methodology of the methods is presented. However, we agree with the reviewer, and repeated it here in the revised manuscript as it will be in the line with
5 the response to another reviewer's comment below (Page 7, Line 30 in the initial version of the manuscript). The following sentence has been added: *As described in section 2.3 the running mean \bar{h} is estimated to remove the large-scale fluctuations in the elevations caused by the geoid used.*

Page 7, Line 29: "Using ... into ..." I don't understand the meaning here.

10 The sentence has been corrected as follows: *Using the JPL versus GSFC scales when applying the LLE method results in freeboard differences of about 4 cm*"

Page 7, Line 30: The authors might want to state why they switched to the EGM08 geoid here.

The following two sentences about selection of the geoid used have been added: *Note that in our analysis of the freeboard
15 retrieval methods we use data from the newer EGM geoid provided with ICESat data in contrast to the ArcGP geoid used by the JPL and GSFC. Although overall effect of geoid selection on freeboard values is small, we observed some improvements after switching to the EGM geoid, which are discussed in this section below.*

**Page 7, Line 31: "from the use of" ... I suggest to write that the authors have obtained these; these are not just used,
20 right?**

The sentence has been changed to: *By applying other combinations of averaging scales we found that these differences depend mainly on h_{s1} value.*

**Page 8, Line 1: "Freeboard ..." The authors could note that positive biases also occur along the coasts - something
25 which also confirms the results of Kern and Spreen (2015).**

In order to mention a bias along the coast the next sentence after this has been changed to: *A positive bias in freeboard estimates when using longer segments for along-track averaging, as well as tendency for enhanced biases along the coast (see Figure 3), are also reported in Kern and Spreen (2015) for the Weddell Sea in Antarctica.*

30 **Page 8, Lines 1-2: "patterns are similar" I don't agree here because, e.g., in the Greenland Sea, the Fram Strait area and the Kara Sea the pattern differ.**

Agree, the sentence has been changed to: *Freeboard differences due to different averaging scales for calculation of \bar{h} are small and their patterns have the features of those related to geoid uncertainty shown in Figures 3b and 3c.*

Page 8, Lines 3-6: Why should we note the roughness here?

We suppose that the roughness increases the uncertainty in determination of sea level references. In order to elucidate it we have changed this sentence to: *Freeboard differences and their variability associated with along-track averaging scales may also be linked with the surface roughness, which increases uncertainty in determination of the sea level references. This is confirmed by looking at the period FM05 when the largest values of σ_{25} are observed (Table 2, line 3 and Figure 6a).*

Page 8, Lines 7-8: The text refers to one map in Figure 2d) but the figure actually shows two maps.

We have added 'left' and 'right' in the text when referring the Figure (Figure 3d in the revised manuscript).

10 Page 8, Line 11: The authors could mark the location of these ridges in the respective map.

We have marked the location of the Gakkel and Lomonosov ridges on Figure 3c using abbreviations 'G' and 'L' respectively, and added their description in the figure caption.

15 Page 9, Lines 7-11: Why are the authors doing this? I guess I missed the motivation for discriminating between FYI and MYI. See also my general comment 2.

In order to explain a rationale for discrimination between FYI and MYI we have change the sentences and their order in this paragraph as follows: *The obtained freeboard differences are small on average, ranging within ± 5 cm, while the presence of significant regional discrepancies should be noted. Since the maps in Figure 4 show a clear distinction between the differences over thin and thick ice for some of the ICESat periods, we estimated differences between freeboards separately for the first-year ice (FYI) and multi-year ice (MYI) regions over the same 25-km grid cells.*

Page 9, Lines 19-21: I don't understand this sentence. Basin wide and MYI seem to contract each other. And if the TP method does not provide (enough) freeboard measurements in a 25 km grid box (or track segment) then there are no data from the TP method, right, and can therefore not be compared with data from the other method?!?

25 We mean that the difference between freeboards would be larger for the means calculated from all grid cells, where freeboard estimates are available over large basin-wide areas (but not only from grid cells, where freeboard estimates available for both datasets). We reformulated the sentence as follows: *One can also expect difference between the basin-wide freeboard means in the area of thick MYI to be even larger when the TP method does not detect any tie-points within some grid cells and, hence, does not provide freeboard estimates.*

30

Page 9, Line 22: (i) can be formulated more clearly perhaps. Isn't the main reason here that the LLE identifies re-frozen leads, where the refreezing has happened some time ago and hence the lead is not covered by thin ice anymore, compared to leads with open water or young ice?

We reformulated this phrase as follows: *(i) the lower-biased estimates in the LLE method due to using the measurements over refrozen leads or ice within the 25-km range for the calculation of local sea level references*

Page 9, Line 23: What does absence of local detection mean? Aren't SSH values interpolated across gaps?

5 Kwok et al. (2008, 2009) fill the gaps after conversion of freeboard to sea ice thickness by interpolation of the gridded estimates, while interpolation of the initial freeboard estimates is not applied.

Page 9, Lines 25-27: The statement that more leads are found over FYI area than in MYI areas could perhaps be confirmed from the available literature: Willmes and Heinemann, Remote Sensing, 2016; Ivanova et al., The Cryosphere, 2016; Röhrs et al. The Cryosphere, 2012; Brohan and Kaleschke, Remote Sensing, 2014.

10

We thank reviewer for providing us with the list of relevant references and added them in this sentence.

Page 10, Line 5: "We checked that our results ..." How did the authors do that? What was the measure used to do this?

15 The sentence has been changed to: *We note that visually the curves corresponding to ON05 and FM06 periods are in agreement with those reported in Kwok et al. (2007).*

Page 10, Line 6: The regression curves mentioned are not shown, right?

We believe that all 10 regression lines presented on two plots on the figure (Figure 5 in the revised manuscript) will be tangled and not separated from one another. Therefore, we presented only two averaged regression lines: one for autumn periods and one for winter periods. Following the other reviewer's comments below we also elaborated the section 3.3.2 in order to make it better structured and containing additional description of the figures.

20

Page 10, Line 7: Why "thick MYI"? Isn't simply "MYI" sufficient here?

25 Agree, it has been changed. Please, note that this sentence has been moved to the second paragraph of this section according to the changes we have made to make the section better structured.

Page 10, Line 10: Contributions of tie points with "larger distances" dominate? Why larger?

As stated in the previous sentence the sea level reference for each 25-km segment is estimated by averaging the h_r values corresponding to the tie-points, weighted exponentially by the distance from the regression model. This weighting by the distance implies that larger distances will dominate.

30

Page 10, Line 11: I am a bit confused here. I thought that sigma_25 is computed for nonoverlapping 25-km segments in the TP method. Did you change this and now use a running 25 km segment?

The σ_{25} is computed in the same way for both the LLE and TP methods, i.e. as a running mean, while sea level references in the TP method are estimated for non-overlapping 25-km segments using selected tie-points.

Page 10, Line 12: Why do tie points with lower h_r contribute more than those with larger h_r . I don't fully understand this.

The tie-points with lower h_r (i.e. more negative values) have larger distance from the regression line (for the same σ_{25}). Since the sea level reference is estimated by averaging the h_r values weighted by the distance from the regression line, the tie-points with lower h_r have a greater contribution.

Page 10, Lines 14-15: The authors write about the missing linear relationship between σ_{25} and freeboard for freeboard values approaching 0 and refer to Figure 6 a) already. I suggest the authors spend more effort to first discuss Figure 4 in more detail and then explain what they do and why, by using Figure 5 and 6. Figure 4 did not yet get too much attention. See also my comments to Figure 4.

We have changed and extended a description of the curves on the Figure (Figure 5 in the revised manuscript) as follows:
However, as shown by the flattening of the curves in Figure 5, the quasi correlation existing between h_r and σ_{25} is lost as it can be seen from the flattening of the curves in Figure 5. Although using a cubic polynomial fit of the data as in Kwok et al. (2007) reduces this flattening, the correlation does not hold towards zero h_r for many ICESat periods. As seen from the averaged regression lines on the Figure 5 a deviation from the linear relationship is more pronounced for the winter periods and starts in a freeboard range from -15 cm to -20 cm. However, we think that this flattening of the curves may not represent an actual and physically-based relationship existing between h_r and σ_{25} . If this is this case, some samples may be unreasonably identified as tie-points in the TP method due to enlarged area below the regression line at h_r close to zero.

Then, we discuss regression lines focusing on the selected ICESat periods, i.e. FM05 and FM08 (shown on the Figure 7 in the revised manuscript). In the revised manuscript, we note, in particular, the motivation for selection of these two particular periods. The following sentences after description of the Figure 5 have been added or changed: *The effect of the flattening of the curves on the result of the regression model is illustrated on the Figure 7 for two winter periods: FM05, when the highest σ_{25} values are seen for h_r close to zero, and FM08, for which the largest discrepancy between LLE and TP results are observed. The computed regression line for the FM05 period reduces the inversed correlation between h_r and σ_{25} , but still deviates from the linear relationship (see red dashed lines in Figure 7a). In FM08 the curve is less noisy but correlation is even inversed with an inflection point around $h_r = -5$ cm. As a consequence, the samples detected as tie-points that have a value of h_r close to zero (i.e. measurements taken over areas covered by thinner ice) may contribute more than the other, leading to an artificial increase of the reference sea level height over given segments.*

In the end of this paragraph we provide explanation of missing linear relationship between σ_{25} and h_r : *The fact that the linear relationship between σ_{25} and h_r does not hold for thin freeboard can be explained by a lower likelihood for samples with $\Delta R > 0.3$ to represent actual leads. This is illustrated by the increase of the standard deviation of σ_{25} and the decrease in number of samples used in evaluating σ_{25} for low absolute values of h_r (Figure 7b and 7c, red). Note that this is*
5 *consistent with the more pronounced flattening obtained for the winter periods, when variability of the surface roughness is larger.*

Page 10, Line 18: "contribute more than the other" Why?

In the inversely distributed part of the relationship between σ_{25} and h_r the distance of tie-points from the regression line for
10 a given σ_{25} becomes larger toward zero h_r . The tie-points with larger distance from the regression line have larger contribution.

Page 10, Lines 28-29: "For example, differences ..." I don't see this.

We have reformulated some sentences in this paragraph for more detailed explanation. Since in this paragraph we discuss the
15 Figure 6 (as in the revised manuscript), we have moved here also one sentence with the description of the Figure 6b from the beginning of section 3.3.2. Therefore the second part of the paragraph has been changes to: *In general, the number of tie-points and roughness are anti-correlated and their spatial patterns match very well with the pattern of multi-year versus first-year ice as shown in Figure 6 for some selected ICESat periods discussed in the text, i.e. FM05, ON05, FM06 and FM08. The number of detected tie-points within each 25-km non-overlapping segments ranges from a few tie-points (i.e. <*
20 *10) over MYI to several tens (i.e. > 25) over FYI (Figure 6b) and shows a significant spatial variation. A surface roughness represented by σ_{25} (Figure 6a) typically do not exceed 10-15 cm over FYI, although it may be locally more than 20 cm as for FM05 period, when the largest roughness is observed (Figure 6a). Therefore, the differences of freeboard estimates in FM05 are primarily related to the surface roughness, and to a lesser degree to number of detected tie-points, which is comparatively low. In contrast, the large number of detected tie points plays a key role in FM08, while surface roughness*
25 *over FYI is low.*

Page 11, Line 5: "15 samples" as are shown in Figure 5?

The figure referred by the reviewer (Figure 6 in the revised manuscript) shows the gridded mean number of tie-points within
25-km non-overlapping segments, while in this sentence we discuss number of samples used to form the curves of
30 relationship between σ_{25} and h_r . These curves are formed for 1-cm bins for h_r and we require that at least 15 samples should be available for each bin.

Page 11, Line 12: Why is this? The authors could not yet convince me. The dark blue curves, for instance, seem to be the least noisy ones, for instance.

In the TP method proposed by Kwok et al. (2007) the tie-points are identified as the points located below the regression line of the relationship between σ_{25} and h_r . As we described in the section 2.3 they found that the area just below this regression line corresponds to the area around the regression line ($\pm 1\sigma$) of the relationship between σ_{25} and the freeboard adjacent to leads identified from SAR images and collocated with ICESat data. For $h_r < -15$ black lines are similar to the red ones, i.e. reproduce the approach by Kwok et al. (2007), while the regression lines derived by applying more stringent selection requirement (cyan and blue lines) are shifted downward compared to the red lines. Thus, selection of tie-points using cyan and blue lines will exclude samples from the area between these lines and black line. Indeed, cyan and blue lines are less noisy. However, this is less important factor because the curves are effectively smoothed when approximation by the cubic polynomial fit is used.

10 In this sentence we have added some clarifications: *From this analysis, using the condition $h_r < \overline{h_{r25}} - 0.5\sigma_{25}$ in this improved TP method appears to be the most appropriate because it corrects the relationships for thin ice and, at the same time, better reproduces the TP algorithm for $h_r < -15$.*

Page 11, Lines 13-19: The correction is about 1 cm (1-2 cm for FYI) ... is this worth the effort? Perhaps a more proper discrimination of MYI and FYI would enhance the results (see general comment 2).

15 As we noted in the reply to the general comment 2 we have tested the suggestion made by the reviewer about discrimination of MYI and FYI using 10% and 90% thresholds, but it does not result in significant enhancement of the results. Since the most remarkable effect from our improvement of the TP method is revealed for the FM08 period, we have added one sentence to emphasis it: *The most remarkable improvement is observed for the FM08 period when the difference of 5-10 cm over vast areas of FYI are reduced to differences ranging within ± 2 cm.*

Page 11, Line 18: "differences remain unchanged" This is not true if one looks at Figure 7.

25 Indeed, for MYI the difference are slightly change, but not significantly. We have corrected the sentence to: *As expected, the differences remain almost unchanged for MYI since our modification of the TP method primarily impacts freeboard estimate over thin ice areas.*

Page 11, general comment to section 3.3.2: I have the feeling that a better description and discussion of Figure 4 would form a better basis onto which this section can root. I have the feeling that this section would be more easy to read once the information is structured a bit better and in chronological order, and by underlining the contribution of the authors more.

30 As noted in our replies to the reviewer's comments above (related to section 3.3.2) and below (related to the figures) we have changed and extended description of the Figures 5, 6 and 7 (as in the revised manuscript), and changed the location of some sentences within the section in order to make it clearer and better structured.

Page 11, Line 24-25: "rather uniformly distributed" Why is this the case? Does this make sense in the eyes of the authors?

A correction for snow depth accumulated in leads as estimated using the function proposed by Kwok and Cunningham (2008) is limited by 5 cm. Due to this limit the correction looks uniformly distributed when using the color scale with the range of +/-15 cm, which we use in the manuscript for the maps of freeboard differences. However closer inspection of the maps shows that there are some spatial variations of correction. Therefore, we have changed and extended the description of correction for snow depth as follows: *The adjustment of freeboard included in the TP algorithm and related to snow depth in refrozen leads is limited by 5 cm and is about +2-3 cm on average for all periods considered (Table 3, first line). Due to the fixed limit, this correction is rather uniformly distributed over the Arctic although we note that the lowest values are observed for thin ice in the Arctic seas, i.e. in the warmer regions with slower initial ice growth in leads (Figure 9a for ICESat periods ON05 and FM06).*

Page 11, Line 29: "from and after ON05" This could be gain related - see general comment 3.

We believe that correction for lead width being estimated as a product of freeboards and factor of $1.1 + 0.1 \left(\frac{R_{snow} - R}{R_{ice}} \right)$ following Kwok et al. (2009), depends primarily on the freeboard values. This is also confirmed by spatial distribution of the correction, which is correlated with freeboard distribution.

Page 12, Lines 1-2: I am wondering whether it would make sense to show these values in Table 2 as well ... or in a new Table 3 (see below).

In the revised manuscript we have added a new table (Table 4 in the revised manuscript) where FYI, MYI and overall results are presented following the reviewer's comment below.

Page 12, Line 3: "and over the ICESat period" What is meant here? I would add an "is" before "ranging".

Agree, it has been corrected.

Page 12, general comment: I am wondering why the authors do not show maps generated with their own "optimal" approach?

We agree that our own estimates also can be presented. However, as we focus on the analysis of the algorithm performance and their comparison, we suggest to give these freeboard maps in the supplementary materials.

Page 12, Lines 8-13: I would find it helpful to have the colors used in Figure 9 to be written in parentheses in the text after the respective variable. Does the variation of valid grid cells have an influence on your results?

Agree, the colors have been written in parentheses in the text.

Selection of the grids available for all considered datasets change the mean values within a range of few centimeters. However, as we noted above, in Table 4 and Figure 10 (as in the revised manuscript) we present the mean values calculated from all grids (i.e. without selection of the collocated grids), but only for the area considered in the JPL product. Therefore, the last sentence in this paragraph have been changed to: *Note that these results correspond to the area considered in the*
5 *JPL product, i.e. to the Arctic Ocean without surrounding Arctic seas such as Greenland Sea, Barents Sea, Kara Sea and Baffin Bay. Correspondingly, the difference of 0.42 m between sea ice thickness in the JPL and GSFC products found by Lindsay and Schweiger (2015) was derived using randomly selected samples over Arctic Basin. Although excluding of the above-mentioned seas does not significantly impacts the difference between the results, the freeboard means are changed within the range of ± 2 cm depending of the ICESat period. The average freeboards obtained from the different methods and*
10 *corresponding to the results shown on Figure 10 are recaped in Table 4.*

Page 12, Line 19: "it may underestimate ..." Why? Where did the authors demonstrate this? A link to the respective figure in this paper would help.

In the revised manuscript we have changed and added the description of the results. In particular, the description of
15 comparison between the freeboards obtained from the TP and LLE method has been extended in the section 3.3.1 and 3.5. Please, find this description below in the response to another reviewer's comment of Page 13, Line 29-32 of the initial version of the manuscript.

Page 12, Line 21: "30 cm" I would read 35 cm from Figure 9.

20 Figure 10 (as in the revised manuscript) shows average freeboard over the whole Arctic, FYI and MYI, while in this sentence we refer only to the lowest part of MYI with freeboard of around 30 cm. In the revised manuscript we rewrote the second and third paragraphs of section 3.5 including the sentence referred by reviewer. This sentence has been changed to: *The lower freeboard obtained by the TP method is mostly observed over FYI and part of MYI in the central Arctic especially in the first three ICESat periods (ON03, FM04 and ON04) (Figures 4 and 10).*

25 With regards to this, we have also changed one sentence in section 3.3.1 (first sentence in the third paragraph) to: *Positive differences are obtained over large areas of FYI and thin part of MYI for most of the ICESat campaigns, with a peak in FM08.*

Page 12, Line 22-23: Are you referring to Figures 4 and 6 here?

30 That is correct. The references to these figures (5 and 7 in the revised manuscript) have been added in this sentence.

Page 12, Line 23: "have greater weight" Why?

As discussed in the sections 2.3 and 3.3.2 sea level references are estimated by averaging the h_r values corresponding to the tie-points and weighted exponentially by the distance from the regression line of the relationship between h_r and σ_{25} . Tie-

points with lower elevations (more negative h_r) have larger distance to the regression line for a given σ_{25} and, therefore, have greater weight.

Page 12, Line 26: neighboring to leads ... ICESat data." I don't understand how this worked out without additional reference data of freeboard; were airborne data used?

As we describe in section 2.3, Kwok et al. (2007) identified leads using SAR images and then looked on the elevations derived from collocated ICESat measurements. Please, note that in the revised manuscript we split this sentence in to two and moved to the last paragraph of section 3.3.1 where we discuss the reasons of the differences between TP and LLE results. Please, find this description below in the response to the reviewer's comment of Page 13, Line 29-32 in the initial version of the manuscript.

Page 12, Line 26-27: "However, the difference between freeboards retrieved ..." Is this the desired goal? It sounds like GSFC is the truth against which your method can and is referenced. Is this the case? What I recommend the authors to do is to also discuss the results of using the GSFC approach (LLE) with 25 km window. Figure 9 illustrates that the "regular" GSFC and the 25 km GSFC differ quite a bit. The authors also mentioned that presumably by using 25 km windowing one can get more realistic results than using 100 km windowing. I feel that this part of the discussion is currently missing in the paper.

We have extended discussion of the differences between the results derived from the LLE method with 25 km and 100 km windows in sections 3.2 and 3.5. The following changes have been made.

In second paragraph of section 3.2 we extended a sentence explaining the difference between freeboards to: *This is expected from the fact that considering a larger window increases the chance to include lower h_r dips in the calculation of h_{sl} , although we use the same fraction of the lowest elevations (1%) for selection of tie-points.*

In fourth paragraph of section 3.2 we have changed and extended motivation for choosing of 25-km scale for comparison of the TP and LLE methods by adding one argument, and mention also a drawback of this shorter scale:

In order to avoid this bias when comparing the LLE and TP methods, we chose to use an averaging window of 25-km to calculate \bar{h} and h_{sl} (as in Kwok et al., 2007) for the three following reasons. First, applying of the TP method using the same scales as in (Kwok et al., 2007) allow us to analyse the performance of the algorithm applied to generate the JPL product. Second, using a smaller window is found to result in reduced dependency of the freeboard on the geoid used in the retrieval process. Indeed, in this case the correction of the geoid for $\bar{h} - h_{sl}$ as well as the fact of using a recent geoid like EGM08 no longer has any impact on freeboard along the ridges in the Arctic, as opposed to what we reported above when using larger spatial averaging. Third reason is that, as demonstrated by Kern and Spreen (2015), a more valid freeboard can be retrieved using the LLE method if the length of along-track segments considered for the selection of the lowest elevations when estimating the local sea level reference h_{sl} is equal to, or less than, the size of the smoothing window used to calculate \bar{h} values. It means that when using 100-km window for determination of h_{sl} the same (or larger) scale would be preferred to

calculate \bar{h} , so that fluctuations of the geoid would be not properly taken into account properly. From the other side, and as we have shown above, using a shorter length for the averaging windows when applying the LLE method results in lower freeboards that can be interpreted as underestimates due to poorer sampling of the sea level references. The TP methods, in contrast to the LLE method, is based on selection of tie-points using a physical relationship, and its performance can be assessed by comparison of the results derived from the TP and LLE approaches.

In section 3.5 we have extended description of freeboard differences to:

Because of the use of different averaging scales to calculate sea level references, the sea ice freeboards we estimated using the LLE method with a 25-km averaging window are lower by ~3 cm on average as compared to those of the GSFC product for all ICESat periods (Figure 10, red and black lines). As we discussed in section 3.2, this can be explained by the increased likelihood for the lowest elevations to correspond to actual sea level height when using a larger window (e.g. 100 km as applied to produce the GSFC dataset).

Page 12, Line 30: The authors please check whether they are indeed referring to "sea ice freeboard".

Indeed, we refer to the total freeboard, but not sea ice freeboard. This has been corrected.

15

Page 12, Lines 30-32: 7 cm ... 6 ... 0.42 m ... these numbers and/or conclusions are contradicting the notion of the "ten times" mentioned in the introduction.

Indeed, in contrast to the introduction, we refer here to the total freeboard, but not sea ice freeboard. This has been corrected.

Page 13, Line 2: "_ 1 cm" I am suggesting that the authors add another table in which they show at one glance the results of this paper in comparison to the other results. I guess it would increase the readability of the paper and would also help the authors as well as the readership to estimate the relevance of the optimizations carried out by the authors.

The Table 4 with the mean freeboards corresponding to those presented on Figure 10 as well as the reference to this table in Summary section has been added.

25

Page 13, Lines 3-7: These lines read as a repetition of information which is also given in the conclusions.

Yes, indeed, we repeat the conclusion from the second sentence of this paragraph in the section Conclusions (the first sentence is actually has not been repeated). However, we believe that it is reasonable to repeat one of the main findings of the paper in the Conclusions.

30

Page 13, Line 13: Please mention what the database and region for which this number is valid for.

We added this information in the first paragraph of section 3.5 where we explain that the results presented on the Figure 10 are related to the Arctic Ocean without surrounding seas. Please, find this description above in the response to the reviewer's comment of Page 12, Lines 8-13 of the initial version of the manuscript.

5 **Page 13, Lines 28-29: "inadequate weighing" To which method does this refer?**

We have specified that this refer to the TP method as follows: ... *their inadequate weighting in these calculations when applying the TP method.*

10 **Page 13, Lines 29-32: This reads a bit strange. The authors improve the TP such that it agrees better with the GSFC_100km - over FYI basically - and then recommend to use the TP. One could ask why not simply use the GSFC? Perhaps the authors can stress more why GSFC potentially has more uncertainties / biases than the TP ... and the authors could comment on GSFC_25 km as well.**

In sections 3.3.1, 3.4 and 3.5 of the revised manuscript we changed and extended description of the differences between freeboards retrieved by the TP and LLE methods and of the effect of adjustments for snow depth in leads and lead width.

15 In the last paragraph of section 3.3.1 we added or changed the following sentences:

The positive differences are most likely reflect the underestimation of freeboard retrieved by the TP method as was found by Kwok et al. (2007) from comparison with the freeboards adjacent to leads detected on satellite images and collocated with ICESat data. They showed that the freeboard underestimation was on average of 1.3 to 4 cm for ON05 and FM06 periods, and explain this by the fact that samples, which are identified as tie-points, do not always represent leads or the thinnest ice.

20 *In order to explain why freeboard differences are observed primarily over FYI areas we investigated the performance of the algorithm used in the TP method.*

In the end of section 3.4 we added the following paragraph:

25 *In principle, the freeboards derived by the LLE method can also be corrected for snow depth in leads and lead width, but this was not done for the GSFC product. Although the LLE method selects only the lowest elevations to determine the local sea level, these samples may be contaminated by snow accumulated in leads or by the neighbouring sea ice surface within the laser altimeter footprint. As shown above from the comparison of freeboards derived by the LLE and TP methods, the LLE method has a weakness mostly over the thickest ice due to lack of leads. The GSFC product, in addition, was derived using longer averaging windows that, in general, increases a likelihood for the lowest elevations to represent a true height of sea level. Therefore, we think that when using the LLE method the empirical functions proposed by Kwok and*
30 *Cunningham (2008) and Kwok et al. (2009) should be modified depending on the averaging scale applied.*

In section 3.5 we added or changed the following sentences:

The sea ice freeboards we estimated using the original TP method are lower by ~3 cm on average as compared to those we obtained with the LLE method with identical 25-km along-track averaging scales (Figure 10, green and red lines). As already shown by Kwok et al. (2007), this study shows that the tie-points for determination of sea level references selected by

the TP method not always represent leads or the thinnest ice. Although the tie-points with lower elevations have greater weight in most cases, the resulting sea surface reference is biased positive, hence leading to lower freeboard estimates (see Figure 5 and 7). Therefore, we suppose that obtained difference between the TP and LLE results reflects the underestimation of the freeboard derived by the TP method. The lower freeboard obtained by the TP method is mostly observed over FYI and part of MYI in the central Arctic especially in the first three ICESat periods (ON03, FM04 and ON04) (Figures 4 and 10). However, the difference between freeboards retrieved by the LLE and TP methods is reduced by more than 30% for the whole Arctic and by 40% for FYI when applying our suggested improvements of the algorithm used in the TP method. At the same time, over thick part of MYI the LLE method tends to give lower freeboards. Although it is not reflected in the mean values on the Figure 10, it can be seen on the maps of the differences (Figure 4 and 8), especially for the FM08 period. This is consistent with our expectations that over thick part of continuous MYI with fewer leads, the use of a relationship between the freeboard and surface roughness for identification of tie-points, as done in the TP method, gives more reliable freeboard estimates.

... Although the samples used for determination of sea level in the LLE method can also be affected by snow accumulation and contaminated by the neighbouring sea ice surface, these corrections were not applied to freeboards in the GSFC product. ... This means that, on average, it is sufficient to account for these corrections by using a larger averaging window and LLE method or that both products underestimate freeboard.

Page 17, Table 1: I suggest to use +/- instead of the "/" to separate mean and standard deviations.

Agree, corrected.

20

Page 18, Table 2: What is the unit? What is the second number in each cell? How do these results relate to your correction? What is meant by "adjustment" here - or in other words: What is the reference? In the second row of the table I would write: "Lead width"

We have added the unit (cm) and noted that that second number is standard deviation in the table caption. We use +/- instead of the "/" to separate mean and standard deviations as in the Table 1.

We are not sure that we understood reviewer's question: "What is meant by "adjustment" here - or in other words: What is the reference?" In the section 2.5. and 3.4 we describe the meaning of these adjustments. In the table caption we also give the references to Kwok and Cunningham (2008) and Kwok et al., (2009) where these adjustments are proposed.

The title of the second row has been corrected to "Lead width".

30

Page 19, Figure 1: - In the maps in b) and c) there is a line artefact in the Fram Strait? The authors could comment on this. - The histograms might benefit from a vertical line denoting zero difference.

We commented this artefact difference in the figure caption as: *An artefact line of negative differences along the 0 longitude is due to an unexplained positive anomaly in the GSFC freeboard estimates.*

Page 21, Figure 2: - The maps in d) lack the annotation which geoid is used.

The annotation for d) has been added.

5 **Page 23, Figure 3 - Why is the difference between LLE and TP on average smaller for FM periods than for ON periods? If the positive elevation bias due to re-frozen leads instead of open leads would be the reason, then I would assume that it is easier to hit an open lead / lead with just very thin ice during ON periods than during FM periods. And therefore I would assume that the effect of leads covered with a finite sea ice thickness causing a bias in the LLE and in turn causing an notable difference between LLE and TP to be larger during FM than during ON. But**
10 **apparently the opposite is the case.**

On this figure (Figure 4 in the revised manuscript) the largest negative freeboard differences over MYI are observed for winter periods, e.g. MA07 and FM08. This is exactly because of fewer open leads in winter. For FYI the positive difference are slightly larger for winter periods. As we discuss in section 3.3.2 of the manuscript, this is related to the fact that flattening of the relationship between σ_{25} and h_r toward $h_r = 0$ is more pronounced for winter periods due to larger surface roughness.

15

Page 25, Figure 4: - The images lacks annotations a) and b) - I would find it helpful to see lines of the regression used in the TP for every ICESat period to understand better the limitations of the TP, to understand whether there are differences between ON and FM periods and learn, where on average, the linear relationship starts to fail. One could do it via dashed lines like in Figure 6. - I would be good to understand the motivation why the authors did chose
20 **the periods shown in Figure 6. - How would these plots look when FYI and MYI were discriminated?**

- The annotations has been added to the figure.

- We have added two average regression lines for autumn and winter periods and discuss them in the section 3.3.2. They show the difference between ON and FM periods and where, on average, the linear relationship starts to fail. A discussion of limitation of the TP method has been extended in the description of the Figures 5 and 7 (as in the revised manuscript), and in
25 the end of section 3.3.1.

- As we noted in the response to the reviewer's comments above we extended and changed description of the Figures 5 and 7 (as in the revised manuscript) in section 3.3.2, where we mentioned the motivation for selection of the periods.

- The measurements over FYI are represented in the part of the curves form 0 to 15-20 cm, while MYI is reflected in the rest part of the curves. Therefore, we believe that curves derived separately from FYI and MYI would be similar to those two
30 parts of the relationship.

Page 26, Figure 5: What is the motivation for showing these periods?

These four periods were selected as the periods discussed in the manuscript in more details. We have mentioned it in the second paragraph of the section 3.3.2 where this figure is discussed.

Page 27, Figure 6: - The caption lacks description of what the dashed lines refer to. -For images c) I suggest to use the same scale at the y-axis and to perhaps also give the total number of valid values used in each image (in c).

- Annotation for the dashed lines has been added.

- 5 - We have also added the total numbers of valid values in c). However, we are not sure that using the same scale at the y-axis is reasonable because in this case it will be hard to see the difference between the curves for the FM08 period.

Page 31, Figure 9: - The thin dashed lines show the results of your optimization. I am wondering why the authors are not highlighting their results more - see general comment 1. - I am wondering whether it makes sense to show the "old" original TP version results when the two corrections mentioned (lead fraction in FoV and snow depth / reflectivity) are needed to come up with "reasonable" results using the TP? One could, instead of showing results from the TP without these two corrections show only the final versions of the TP. - The LLE does not require any of these two corrections, right?

- 15 - As we noted in the answer to the general comment 1 we tried to highlight our results in the context of the difficulties of deriving sea ice freeboard and thickness from satellite laser altimetry. Also we pointed out significant improvement for the FM08 period.

- 20 - In principle, the freeboards derived by the LLE method also can be corrected for snow depth in leads and lead width, though it has not been done in the GSFC product. Although the LLE method selects only the lowest elevations for the tie-points these samples also may be contaminated by snow accumulated in leads or by the neighbouring sea ice surface within the laser altimeter footprint. Therefore, we think that it is reasonable to show the results derived by the LLE and TP methods without these corrections in order to show the difference caused solely by the method of sea level reference determination.

We have added discussion of the effect of these two corrections in the revised manuscript. Please, find this description above in the response to the reviewer's comment of Page 13, Line 29-32 in the initial version of the manuscript.

- 25 **Typos: Page 1, Line 11: ICESat -> ICESat**

Agree, it has been corrected.

Page 3, Line 15: denoted -> denote

Agree, it has been corrected.

30

Page 4, Line 28: accordingly -> according

Agree, it has been corrected.

Page 5, Line 7: Kwok et al., -> Kwok et al.

Agree, it has been corrected.

Page 12, Line 25: "in average" -> "on average"

5 Agree, it has been corrected. Please, note that the revised manuscript this sentence has been moved to the last paragraph of section 3.3.1 where we discuss the reasons of the differences between TP and LLE results.

Page 13, Line 20 & 27: "in average" -> "on average"

Agree, it has been corrected.

10 **Page 14, Lines 3-4: "in average" -> "on average"**

Agree, it has been corrected.

Page 14, Line 17: "6(69)" -> "56(69)"

Agree, it has been corrected.

15

Page 14, Line 19: Please add "-Oceans" in the Journal title.

We are not sure that this comment is correct. Paper by Kurtz and Markus ("Satellite observations of Antarctic sea ice thickness and volume") is published in the Journal of Geophysical Research, but not Journal of Geophysical Research Ocean.

20

2 Response to anonymous referee #2

General comments

I feel the paper has merit for publication, but the presentation can be improved.

25 We have substantially changed and extended description of the results and figures in section 3 in the revised manuscript, in particular, by accommodating the reviewers' comments.

Specific comments

30 **(1) in the abstract, you said "LLE give significantly lower estimate over thick multi-year ice and larger estimates over thin first year ice as compared with the TP". However, from figure 9, it seems in all cases, the freeboard from LLE (Goddard) already higher than from TP (Tiepoints). Please explain.**

In this sentence we note that these differences are derived when applying the same along-track averaging scales, i.e. we refer the results denoted as 'Tiepoints' and 'Lowest level elevation, 25 km' on the Figure 9. The GSFC product is obtained using

longer along-track averaging scale that increases freeboard estimates. We have rewrote most part of the abstract to present our findings in the paper more clear.

5 **(2) in the abstract, line 20-24, you seem say the LLE and TP methods give similar freeboard estimates (that's why you said the ice thickness difference might be due to different other parameters used). However, from the comment (1) above, it is clear that freeboard from LLE and TP are not the similar. Please explain this statement or rewrite it.**

We found that LLE and TP methods give on average similar freeboard estimates if we apply the settings in the freeboard retrieval algorithms as were used to generate JPL and GSFC products. The algorithms applied to produce these two datasets have three differences: 1) method for determination of the sea level references (LLE and TP methods), 2) along-track
10 averaging scale (longer for GSFC product and shorter for JPL product), 3) corrections for snow depth and lead width was applied only in the JPL product. We assess the effect from all these factors and conclude that on average they should roughly compensate each other. The differences noted in the sentence discussed in the comment (1) describe the effect of the difference '1)', i.e. difference of the method for determination of the sea level references. We have rewrote most part of the abstract to present our findings in the paper more clear.

15

Page 2, Line 24, references for using the LLE method, there is important reference missed and suggest to add in here. Xie, H., A. Tekeli, S. Ackley, D. Yi, and J. Zwally, 2013. Sea ice thickness estimations from ICESat Altimetry over the Bellingshausen and Amundsen Seas, 2003-2009, Journal of Geophysical Research, doi:10.1002/jgrc.20179

In the revised manuscript we have included this reference in the list of relevant papers.

20

Page 3, line 1, "JPL estimates" to "JPL estimates of ice thickness"

We have reformulated this paragraph. The sentence referred by the reviewer has been combined with the first sentence of this paragraph and changed to: *The Arctic sea ice thickness from two available products derived from ICESat data by Jet Propulsion Laboratory (JPL) using the TP method (<http://rkwok.jpl.nasa.gov/icesat>) and by Goddard Space Flight Center
25 (GSFC) using the LLE method (Yi and Zwally, 2009) were found to be different by 0.42 m (Lindsay and Schweiger, 2015).*

P4-L20, lowest 1

Agree, this phrase has been changed to: ... *lowest 1% of the h_r values...*

30 **P5-L2, is very confusing, should remove the 25.**

We are not sure that removing '25-km' is reasonable here. Kwok et al. (2007) could, in principle, use another window size to estimate σ_{25} . In addition it explains a subscript for σ_{25} .

P6-L7, "ice" should be "ocean water"?

When defining the variables in this scaling factor, Kwok et al. (2009) describe R_{ow} as reflectivity of bare ice. In order to avoid inconsistency we have changed this subscript to 'ice', i.e. to R_{ice} .

P6-L10-19, it seems part of the method and should be move to the method. If you can make a kind of work flow chart, it would be even better.

In the section 2 we describe the methods of freeboard retrieval that we reproduce, while in this paragraph we outline the results of freeboard comparison. Therefore, we believe that it is more reasonable to keep this part as the introduction of section with the results.

We have made a flow chart of the methods and settings implemented in our study as a new Figure 1. Please, note that due to adding of new Figure 1 and Table 1 in the revised manuscript the numbering of figures and tables is shifted.

Section 3.5 summary should be combined into the conclusion section

We agree that, in principal, this section could be combined with the conclusions. However, the Summary section presents not only overview of the results obtained in this study, but we also discuss the mean time series on the (Figure 10 in the revised manuscript) and mean freeboards in the new Table 4 that have not been presented before. Therefore, we think that this summary is more relevant to the section with the results.

P14-line 7-8, “we obtain similar: : .:”, however, figure 9 shows that they are not similar, although the patterns are similar.

As we noted in the reply to comments (1) and (2) the results become on average similar when we apply the settings in the freeboard retrieval algorithms as were used to generate JPL and GSFC products. The results on the Figure 10 (as in the revised manuscript) denoted as ‘Tiepoints’ (green) and ‘Tiepoints modified’ (green dashed) are obtained without corrections for snow depth in leads and lead width. After applying these corrections the results (blue line) become close to the GSFC product (black).

Some table and figure captions include a lot of information but a kind of strange way as compared with normal captions of figures and tables. I hope they can make changes and improve them. so readers can easily understand the table or figure. In their current format, I have difficult to read them, for example, Table 1, Figure 2 ((d) is not mentioned in the caption and no unit for any of them)

The annotation for d) on the Figure 3 (as in the revised manuscript) has been added. Indeed, the captions to the Table 2 and Figures 2 and 3 (as in the revised manuscript) contain a lot of information. However, on these figures and table we a various sort of information on different plots (in the figures) and lines (in the Table 2). Therefore, it seems difficult to shorten this information without loosing details that are needed to understand their content.

On retrieving sea ice freeboard from ICESat laser altimeter

Kirill Khvorostovsky¹, Pierre Rampal¹

¹Nansen Environmental and Remote Sensing Center, 5006 Bergen, Norway

Correspondence to: K. Khvorostovsky (kirill.khvorostovsky@nersc.no)

5 **Abstract.** Sea ice freeboard derived from satellite altimetry is the basis for estimation of sea ice thickness using the assumption of hydrostatic equilibrium. High accuracy of altimeter measurements and freeboard retrieval procedure are therefore required. As of today, two approaches for estimation of the freeboard using laser altimeter measurements from Ice, Cloud, and land Elevation Satellite (ICESat), referred to as tie-points (TP) and lowest-level elevation (LLE) methods, have been developed and applied in different studies. We reproduced these methods [for the ICESat observation periods \(2003-](#)
10 [2008\)](#) in order to assess and analyze the sources of differences found in the retrieved freeboard and corresponding thickness estimates of the Arctic sea ice as produced by the Jet Propulsion Laboratory (JPL) and Goddard Space Flight Center (GSFC). [Three main factors are found to affect the freeboard differences when applying these methods: 1\) the approach used for calculation of the local sea level references in leads \(TP or LLE methods\), 2\) the along-track averaging scales used for this calculation, and 3\) the corrections for lead width relative to the ICESat footprint and for snow depth accumulated in refrozen](#)
15 [leads. The LLE method and longer along-track averaging scale of 100 km, as used to produce the GSFC datasets, both give larger freeboard estimates comparing to those derived by applying the TP method and 25-km averaging scale as used for the JPL product. These two factors contribute to the freeboard differences in approximately equal proportions, and their combined effect is, on average, about 6-7 cm. The effect of using different methods varies spatially: the LLE method tends to give lower freeboards \(by up to 15 cm\) over the thickest part of multi-year ice and higher freeboards \(by up to 10 cm\) over](#)
20 [first-year ice and thin part of multi-year ice. However, we show that the difference over areas of thin ice can be reduced to less than 2 cm when using the improved TP method proposed in this paper. The corrections for snow depth in leads and lead width are applied only for the JPL product and increase the freeboard estimates by about 7 cm, on average. For the ICESat observation periods \(2003-2008\) it is found that when applying the same along-track averaging scales in the two methods to calculate the local sea level references the LLE method gives significantly lower \(by up to 15 cm\) sea ice freeboard estimates](#)
25 [over thick multi-year ice areas, but significantly larger estimates \(by 3-5 cm in average and locally up to about 10 cm\) over thin first-year ice areas, as compared to the TP method. However, we show that the difference over first-year ice areas can be reduced to less than 2 cm when using the improved TP method proposed in this paper. About 4 cm of the difference in the JPL and GSFC freeboard estimates can be attributed to the different along-track averaging scales used to calculate the local sea level references. We show that the effect of applying corrections for lead width relative to the ICESat footprint, and for snow depth accumulated in refrozen leads \(as it is done for the last release of the JPL product\), is very large and increase freeboard estimates by about 7 cm. Thus, the](#)
30 [different approaches to calculate sea surface references and](#) different along-track averaging scales ~~and approaches to calculate sea surface references~~, from one side, and the freeboard ~~adjustments~~

[corrections](#) as applied ~~in the TP method used to~~when producing the JPL dataset, from the other side, are roughly compensating each other with respect to freeboard estimation. Therefore, [one may conclude that](#) the difference in the mean sea ice thickness ~~found~~ between the JPL and GSFC datasets [reported in previous studies](#) should be attributed [mostly](#) to different parameters used in the freeboard-to-thickness conversion.

5 1 Introduction

The observed thinning of the Arctic sea ice cover during the last 15 years is one of the most sensitive indicators of the climate change (e.g. Stocker et al, 2013; Laxon et al, 2013). The main data source for retrieving the sea ice thickness over large-scale basins is the radar and laser satellite altimeter measurements of the sea ice freeboard, which is used to convert freeboard to thickness assuming the hydrostatic equilibrium of floating ice (e.g. Kwok et al, 2009; Laxon et al, 2003; Laxon et al, 2013; Ricker et al, 2014; Tilling et al., 2015). Using this particular conversion method, the uncertainty of the obtained sea ice thickness is equal to [approximately](#) ten times the one associated with the [ice](#) freeboard estimate. This stresses the need for very accurate altimeter measurements and freeboard retrieval procedure in order to minimize sea ice thickness uncertainty (e.g. [Zygmuntovska et al, 2014](#)), and increase the confidence level associated to the negative trend in Arctic sea ice volume reported in the last 2013 IPCC report (Vaughan et al, 2013).

In this paper we focus on uncertainty of total (snow plus sea ice) freeboard retrieval using laser altimeter measurements from Ice, Cloud, and land Elevation Satellite (ICESat). As compared to satellite radar altimetry, ICESat provides higher accuracy in elevation measurements over a comparatively smaller footprint of ~70 m, with a precision of about 2 cm (Kwok et al., 2004) [and the single-shot accuracy of 13.8 cm \(Zwally et al., 2002\)](#). A key step in freeboard estimation process is the determination of the local sea surface height that is used as reference elevation. The determination of the local sea surface height from geoid, modelled tides, and atmospheric pressure loading is rather uncertain. Therefore, a common method is to calculate a local reference elevation from ICESat measurements over open water (or thin ice) within leads. Several methods to detect such samples were proposed in a number of studies (Kwok et al., 2007; Zwally et al, 2008; Farrell et al., 2009).

The approach proposed by Kwok et al. (2007), referred as the tie-points (hereafter TP) method, is based on identification of sea surface reference points (tie-points), for which the deviation between the measured elevation and the local mean surface exceeds a given value. Indeed, they found a relationship between surface roughness and freeboard adjacent to new lead/crack openings associated with low reflectivity, and therefore used it for tie-points detection. Later, the TP method has been further developed by Kwok and Cunningham (2008) and Kwok et al. (2009), and applied by Kurtz et al (2009) and Kurtz et al (2011) to study sea ice thickness in the Arctic. A similar approach, based on the same roughness/freeboard relationship, has been defined and used by Markus et al (2011) and Kurtz and Markus (2012) to retrieve freeboard of the Antarctic sea ice. However, one should note that the TP method has some limitations as it is based on an empirical relationship that may not be valid for some specific time and location.

Another approach, the so-called lowest-level elevation (hereafter LLE) method, was originally described and used in a study by Zwally et al. (2008) and later applied by, for example, Yi et al. (2011), [Xie et al. \(2013\)](#) or Kern and Spreen (2015) to retrieve freeboard of Antarctic sea ice, and by Yi and Zwally (2009) for the Arctic sea ice. The LLE method is based on selecting a certain percentage of the lowest elevation measurements within the along-track section surrounding every ICESat sample, and assumes that their mean [\(as in \(Yi and Zwally, 2009\)\)](#) or their polynomial fit [\(as in \(Spreen et al., 2006\) and \(Kern and Spreen, 2015\)\)](#) represents the local sea surface height for the given sample. A main limitation of the LLE method is that in case of absence of leads or cracks in the vicinity of a given measurements or if the selected percentage of lowest elevations is larger than the actual number of measurements over leads/cracks, the level of sea surface is overestimated, and consequently the freeboard is underestimated.

~~The Arctic sea ice thickness from two Comparison of available thickness products derived from ICESat data by Jet Propulsion Laboratory (JPL) using the TP method (<http://rkwok.jpl.nasa.gov/icesat>) and by Goddard Space Flight Center (GSFC) using the LLE method (Yi and Zwally, 2009) shows that were found to be different by 0.42 m (part of the difference between the estimates can be attributed to the uncertainty of the freeboard to thickness conversion and in particular to the chosen value for ice density (Zygmuntovska et al., 2014). Lindsay and Schweiger, (2015). found that JPL estimates are thicker on average than those in the GSFC product by 0.42 m, and This difference proposed that it can be caused by the different techniques for determining the local sea level in the freeboard retrieval algorithm, and/or by the different methods in estimating snow depth that is used when calculating ice thickness, i.e. by the uncertainty of the freeboard-to-thickness conversion (Kwok and Cunningham, 2008; Zygmuntovska et al., 2014).~~

In this paper we reproduce the two approaches used to retrieve Arctic sea ice freeboard, i.e. using the TP and LLE methods.

We analyze why these methods lead to differences in local freeboard estimates, show how they are distributed in space and over the ICESat period (2003-2008), and propose an improvement in the freeboard retrieval algorithm used in the TP method. The TP method presented originally in Kwok et al. (2007) was further developed and improved to take into account snow that is accumulated on thin ice in leads (Kwok and Cunningham, 2008) and size of leads with respect to the size of the ICESat altimeter footprint (Kwok et al., 2009). [These two corrections were taken into account in the JPL product. Therefore we quantified The their effect of these two corrections](#) on freeboard estimates, and hence on the difference between the corresponding sea ice thickness products, ~~is also quantified.~~

2 Data and methods

2.1 Data

In this study we use the ICESat level 2 data of Release 32 from 10 laser campaigns, corresponding to periods of ~35 days in autumn and winter named as 2a , 2b , 3a , 3b , 3d , 3e, 3g, 3h, 3i and 3j in the ICESat dataset and that we will hereafter denote with respect to the period covered as ON03-~~(i.e. October November 2003)~~, FM04, ON04, FM05, ON05, FM06, ON06, MA07, ON07, FM08. [The abbreviations ON, FM and MA mean October-November, February-March and March-](#)

[April respectively, followed by the year \(i.e. 2003 to 2008\)](#). We also use along-track freeboard ~~product estimates~~ derived from ICESat (Yi and Zwally, 2009) and available for download on the NSIDC server (<http://nsidc.org/data/nsidc-0393>).

In addition, we use the Arctic-wide multiyear ice fraction dataset used in Zygmuntowska et al. (2014) that was produced by reprocessing the QuikSCAT satellite scatterometer data. [Zygmuntowska et al. \(2014\) used daily averaged gridded \(22.5 km\) data of radar backscatter processed by Brigham Young University \(ftp://ftp.scp.byu.edu/data/qscat/SigBrw\), and converted them into multiyear ice fraction following using](#) the method described in Kwok (2004), ~~in order to get daily estimates~~. We also use the NSIDC, daily, 25-km-resolution sea ice concentration product based on Advanced Microwave Scanning Radiometer (AMSR-E) satellite measurements [\(Cavalieri et al., 2014\) and available for download at https://nsidc.org/data/AE_SI25/versions/3](#).

10 2.2 ICESat data filtering and corrections

Unreliable ICESat elevation estimates were filtered out using waveform parameters of the altimeter returns provided together with ICESat data. [When reproducing the TP and LLE methods, we](#) used the [same filtering criteria that are](#) applied in Kwok et al. (2007), ~~when reproducing both methods~~ in order to compare the algorithms avoiding biases associated with different filtering. We discarded measurements where the receiver gain used for indicating forward scattering in the atmosphere (i_gval_rcv) was more than 30 and the standard deviation of the difference between received ICESat echo waveforms and the Gaussian fit ($i_SeaIceVar$) was more than 60. Saturated waveforms occurred over bright smooth flat surfaces with reflectivity ($i_reflctUC$: ratio between received and transmitted energy) >1 were removed. In addition, we filtered out highly saturated returns with amplitude greater than the saturation index threshold for more than five consecutive waveform gates ($i_satNdx > 5$). The influence of the filtering criteria on freeboard estimates will be illustrated below in an example where we apply a different threshold for the receiver gain parameter (80, as used by Yi and Zwally, 2009).

Then we determine the elevation, h , above the EGM08 geoid (Pavlis et al., 2012) provided with the ICESat data, and apply saturation correction to the measurements with moderately saturated waveforms using the corresponding parameter ($i_satElevCorr$) and flag ($i_satCorrFlg$), as well as the inverse barometer correction to allow for atmospheric pressure loading (Kwok et al., 2006, Zwally et al, 2008). We discard the areas with open ocean, which we define here as the region covered by less than 30% sea ice coverage, according to the AMSR-E ice concentration product from NSIDC.

It should be noted that Kwok et al. (2007) and Yi and Zwally (2009) used ICESat data from the earlier Release 28 and estimated elevations, h , above the ArcGP geoid (http://earth-info.nga.mil/GandG/wgs84/agp/hist_agp.html) rather than the EGM08 geoid. Furthermore, when applying the LLE method Yi and Zwally (2009) first calculated the improved geoid before using it for freeboard estimation. Although freeboard retrieval from satellite altimetry is primarily based on estimation of local sea level in leads [we evaluate](#) the effect of using a different geoid on the results ~~is evaluated~~ in the section 3.2.

2.3 Algorithms used in the TP and LLE methods

In order to remove longer wavelength and large amplitude variability due to geoid, atmospheric loading and tidal errors, the first step for both (TP and LLE) freeboard retrieval algorithms is to determine relative elevations, h_r , defined as the difference between elevations h and their 25-km (like in Kwok et al., (2007)) or 50-km (like in Yi and Zwally, (2009))

5 running means, \bar{h} , as $h_r = h - \bar{h}$ (Table 1). [We evaluate and discuss](#) ~~the~~ the effect of the different scale of spatial smoothing when calculating \bar{h} ~~is evaluated and discussed~~ in section 3.2.

The principal distinction between the algorithms, as noted above, is the difference in the method used to determine the sea surface references. For the LLE method, Yi and Zwally (2009) assumed that the lowest 1% of the measurements along the satellite track represent elevations over open leads and therefore can be used for estimation of the local sea level. Hence, they

10 determine elevations of sea level, h_{sl} , as the mean of the ~~1%~~ lowest [1% of the](#) h_r values within ± 50 km around each measurement point. Therefore, in this approach the number of points used for the determination of sea level depends only on measurements' availability. ~~Since~~ ~~the~~ The distance between ICESat samples along track is 172 m, ~~and if~~ we assume that all samples are reliable, 6 of about ~~600-580~~ measurements within the 100-km range are used to calculate the local sea surface level. However, in case there is no open water within the 100-km range, the calculated h_{sl} will be the height of thin ice

15 rather than the sea level height, leading to an underestimation of the sea ice freeboard.

For the TP method, Kwok et al. (2007) determined sea level from ICESat samples, tie-points, identified according~~ly~~ to specified requirements. From the analysis of [SAR images from RADARSAT](#) satellite ~~images~~, Kwok et al. (2007) found a linear relationship between along-track elevation variability and freeboard values adjacent to new openings at the same locations where leads were identified and collocated with ICESat data. Although determination of this relationship provides

20 a tool for detection of tie-points, the procedure of visual inspection of satellite images is time-consuming, and can be applied only for regional analysis (e.g. Markus et al., 2011). Therefore Kwok et al. (2007) proposed to use the relationship between elevation variability and negative h_r with dips in reflectivity, which are found to be associated with young ice in leads. Elevation variability is defined as the standard deviation σ_{25} of the detrended h_r within a 25-km running window. The dips in reflectivity are defined as when the difference between the local reflectivity for a given sample and the background

25 reflectivity, ΔR , is larger than 0.3. The background reflectivity is estimated as the average reflectivity of ~~all~~ the measurements within 25 km around a given sample that are greater than $\bar{R} - 1.5\sigma$ [with \$\bar{R}\$ being the mean reflectivity of all the measurements within 25 km around that given sample](#). Then they select all samples corresponding to the points below

the line obtained from a regression model, here a cubic polynomial, to be used as tie-points. That is, they select samples for which σ_{25} is less than that determined from the cubic polynomial model for the given h_r . Kwok et al., (2007) identified two

30 sets of tie-points: one set consists of samples located below the regression model which also have $\Delta R > 0.3$, and the other one includes all the samples corresponding to the points located below the regression model without constraining the value of ΔR . They found good agreement between ~~these~~ two sets of tie-points as well as ~~their~~ agreement [of these sets](#) with high-quality tie-points determined from collocation with satellite images. Since [a sampling density](#) ~~the number~~ of tie-points in the

former set is not sufficient for ~~the~~ basin-wide studies, both sets of detected tie-points are used for the calculation of sea level. The sea surface references are estimated for 25-km non-overlapping segments as an average of h_r values corresponding to the tie-points, weighted as the exponential function of the distance to the line obtained from the regression model. The higher σ_{25} , which is characteristic of the surface roughness ~~around-of a~~ given sample, the lower is the h_r ~~is required to be for this~~
5 ~~sample~~ to qualify this sample as a tie-point. This weighting method of the tie-points utilizes their likelihood to be a reference point, and is particularly important when many tie-points are detected within a 25-km segment. Since the position of the line issued with the regression model varies over seasons and years, Kwok et al., (2007) proposed to apply the same regression model to each ~~separately for~~ each ICESat observation period. We discuss the influence of the ~~chosen~~ regression model ~~is~~
~~discussed~~ in ~~the~~ section 3.3.

10 2.4 Correction of geoid

In Zwally et al. (2008) and Kwok et al. (2007) a difference between running mean elevations, \bar{h} , and the determined elevation of sea level, h_{sl} , was found in order to characterize the unresolved residuals in the sea surface height. Since the spatial pattern of the differences is found to be consistent for different ICESat campaigns these residuals were mainly associated with the characteristics of uncertainties in the static geoid, and to a less degree as coming from time-varying
15 components or noise in the freeboard estimation process. Therefore Yi and Zwally (2009) applied this difference, $\bar{h} - h_{sl}$, to correct geoid heights, and used this new improved geoid for retrieving the freeboard. We determined the differences $\bar{h} - h_{sl}$ for each along track measurement, and examined the effect of this geoid adjustment on freeboard estimates in section 3.2.

2.5 Adjustments for snow depth and area of sea surface references

After being presented in Kwok et al. (2007) the TP method for freeboard retrieval was further developed by implementing
20 two corrections of sea surface references based on functions determined empirically. One correction is an adjustment of the elevations at the location of ~~in~~ tie-points for taking into account the depth of snow accumulated over young ice (Kwok and Cunningham, 2008). This adjustment is based on the contrast difference in reflectivity existing between sea ice and snow surfaces, and is estimated as a function of $(1 - \Delta R)$. The correction of the sea surface references varies within the range 0 to 5 cm (Fig 2b in Kwok and Cunningham (2008)). It can be noted that increase of reflectivity over young ice may also reflect
25 the effect of frost flowers growth. Therefore, the function of reflectivity that accounts for accumulated snow and used to determine the correction may be different in the presence of frost flowers.

Another correction accounts for the fact that ICESat measurements over tie-points are contaminated by the neighbouring sea ice surface within the laser altimeter footprint. In Kwok et al. (2009), it was proposed to multiply all the freeboard measurements by a factor of $1.1 + 0.1 \left(\frac{R_{snow} - R}{R_{owice}} \right)$, where $R_{snow} = 0.7$ and $R_{owice} = 0.25$ are the typical reflectivity of snow
30 and ice respectively, and R is the reflectivity of the ICESat measurements. This correction increases with freeboard height and decreases with the reflectivity in ICESat samples.

3 Results and discussion

In this section we compare different freeboard estimates retrieved using our implementations of the TP and LLE methods [as illustrated on the flow chart in Figure 1](#). First, we test the agreement between freeboards obtained using our implementation of the LLE method and those provided in the GSFC product. Then, we analyze how the choice of different along-track averaging scales and geoid definition affect the freeboard estimates when using the TP and LLE methods, and therefore how it can partly explain the differences found between the JPL and GSFC products. We also quantify the effect of applying different approaches for determination of sea surface references in the TP and LLE methods when choosing the settings, which give consistent freeboard retrievals. Based on these analyses we propose an improvement of the freeboard retrieval algorithm used in the TP method. We also estimate the effect on freeboard estimates of applying corrections accounting for snow depth in tie-point areas and for the size of leads, as it was done for ~~the last release of~~ the JPL product. Finally, we proceed with a comparison of the obtained freeboard when using the different methods and parameters, and we summarize our findings.

3.1 Comparison of GSFC product with freeboard retrieved using the LLE method

We checked consistency between the freeboards retrieved in this study and those available from the GSFC product by following the LLE method described in Yi and Zwally (2009): the elevations were used relative to the ArcGP geoid corrected for $\bar{h} - h_{sl}$ residuals, the calculation of \bar{h} values was made using an along-track smoothing window of 50-km, and the 1% lowest elevation measurements over the 100-km along-track segment centered on each sample were used for estimation of the reference sea level h_{sl} . We compared our results with the freeboards of the GSFC product by first computing the differences between the freeboards calculated along track for each sample before computing their averages over a regular 25-km grid covering the data domain. Maps of freeboard estimates as well as maps of differences and their distribution for the ON05 and FM06 ICESat periods are presented in Figure [2+1](#). The mean and standard deviation of the differences for the other periods are recapped in Table [2+](#) (first line). The mean differences are small, i.e. around ± 2 cm, indicating good agreement between the estimates. The remaining discrepancies can be attributed to different data filtering, and [partly-possibly](#) to the differences existing between data releases ([e.g. improvements in saturation correction](#)).

In particular, persistent underestimation of the freeboard thinner than 15 cm by up to 10 cm for the ON05 and FM06 periods (Figure [2+b](#) and [2+d](#)) can be explained by different threshold values used for the receiver gain parameter. Indeed, following Kwok et al. (2007) we used ICESat measurements with ~~the~~ gain values of less than 30, while Yi and Zwally (2009) chose to set this threshold value to 80, thereby involving more data in their analysis. This additional portion of the data is more affected by atmospheric forward scattering, leading to measurements showing a larger range and shifted towards lower elevation values. For thin ice the likelihood is high for these elevations to be lower than the neighbouring ones along the track, and as a consequence for them to be used for determination of sea surface reference, which may finally result in higher freeboard estimates. We checked that applying the exact same threshold as in Yi and Zwally (2009) for the receiver gain

parameter increases the agreement between the estimates, as one can see from the removal of the very negative (in blue) differences present over the Kara, Laptev and Chukchi Seas in the maps of Figure 2+c as compared to the maps of Figure 2+b.

One should note however that ~~in addition~~ Yi and Zwally (2009) also used a pulse-broadening parameter for the data filtering that is not applied here. This parameter primarily depends on the width of the echo waveform and, among other effects, accounts also in part for atmospheric forward scattering. This explains why we obtain a noticeably higher freeboard as compared to the GSFC product for some ICESat periods like e.g. ON03 and ON04 when setting the threshold value for receiver gain to 80 (Table 2+, second line). Thus we think that setting the threshold value for receiver gain to 30 is roughly equivalent to the filtering settings applied by Yi and Zwally (2009) to account for forward scattering, but has the advantage of being more efficient over regions of thin ice. It should be noted that Yi et al. (2011) used different thresholds for gain in order to account for the reduce of gain with the age of the ICESat's lasers due to decrease of the transmitted power. However, in this study, we compare freeboard estimates with the GSFC product, which is derived by Yi and Zwally (2009) using constant setting for the gain threshold. The largest biases and their variability observed in the periods ON03 and ON04 correspond to the first operation periods of laser 2 (campaign 2a) and laser 3 (campaign 3a) respectively. Therefore, one may presume that the pulse-broadening parameter applied by Yi and Zwally (2009) is also affected by the instability in the power transmitted by the ICESat's lasers.

3.2 Sensitivity of freeboard estimates to LLE method parameters and geoid definition

The JPL and GSFC products, as noted above, are generated using different along-track averaging scales to calculate \bar{h} values, i.e. 25 and 50-km respectively. The length of the along-track segment used for estimation of the local sea level references h_{sl} is also different: 1% of lowest elevations available over 100-km around each sample for the GSFC product (Yi and Zwally, 2009), while 25-km non-overlapping segments are used for the JPL product (Kwok et al., 2007). Therefore, before comparing the freeboard retrievals obtained with the two methods we checked, as an example, how the choice of different averaging scales influences the results ~~when using of~~ the LLE method. Although both freeboard retrieval algorithms are based on the difference between ICESat elevations over sea ice and the neighbouring leads, which makes them almost fully independent from the geoid accuracy, we also looked at how the choice of geoid influences the results.

We compared the freeboards retrieved using the LLE method with a window's size of 25-km and 50-km to compute \bar{h} , and using along-track averaging segments of 25-km and 100-km to estimate h_{sl} . Note that these averaging scales correspond to those used to produce the JPL and GSFC freeboard estimates. As described in section 2.3 the running mean \bar{h} is estimated to remove the large-scale fluctuations in the elevations caused by the geoid used. Note that in our analysis of the freeboard retrieval methods we use data from the newer EGM geoid provided with ICESat data in contrast to the ArcGP geoid used by the JPL and GSFC. Although overall effect of geoid selection on freeboard values is small, we observe some improvements after switching to the EGM geoid, which are discussed in this section below. Using the JPL versus GSFC scales ~~into when applying~~ the LLE method results in freeboard differences of about 4 cm (Table 2+, line 3 and Figure 3+2a for the FM06

period). By applying other combinations of averaging scales we found that these differences ~~are actually mainly coming from the use of different~~ depend mainly on h_{sl} values. This is expected from the fact that considering a larger window increases the chance to include lower h_r dips in the calculation of h_{sl} , although we use the same fraction of the lowest elevations (1%) for selection of tie-points. Freeboard differences due to different averaging scales for calculation of \bar{h} are small and their patterns ~~are similar to~~ have the features of those related to geoid uncertainty shown in Figures 32b and 32c. A positive bias in freeboard estimates when using longer segments for along-track averaging, as well as tendency for enhanced biases along the coast (see Figure 3), ~~is~~ are also reported in Kern and Spreen (2015) for the Weddell Sea in Antarctica. ~~One should also note that~~ Freeboard differences and their variability associated with along-track averaging scales may also be linked with ~~are dependent on~~ the surface roughness, which increases uncertainty in determination of the sea level references. This is confirmed by looking at the period FM05 when the largest values of σ_{25} are observed (Table 24, line 3 and Figure 65a).

The residuals $\bar{h} - h_{sl}$ applied for the correction of the ArcGP geoid are calculated using the same settings as in Yi and Zwally (2009) and shown in Figure 32d (left) for the period FM06. The spatial distribution of the residual is similar for the other periods and is in agreement with those obtained by Kwok et al (2007). The effect on freeboard estimates is small, ranging within ± 2 cm, over most of the Arctic basin (Figure 32b). The only noticeable effect on freeboard is found in the areas of the Gakkel and Lomonosov ridges, where freeboard is reduced by about 5 cm. The areas of positive differences along the East Greenland and Canadian Arctic coasts correspond to regions of largest freeboard (Figure 24a), which is itself correlated with local surface roughness (Figure 55a), more than to the distribution of the geoid correction. The $\bar{h} - h_{sl}$ adjustment of EGM08 geoid is proportionally lower everywhere in the Arctic by about 13 cm, which corresponds to the higher level of the EGM08 geoid (Figure 32d, right). The effect of adjustment of the EGM08 geoid for $\bar{h} - h_{sl}$ on the freeboard is of the same order in means (Table 24, line 4) and even less evident along the ridges in the central Arctic, which likely results from the overall better quality of this more recent geoid, and in particular from the better representation of small scale features. Small freeboard differences are also obtained when using the EGM08 instead of the ArcGP geoid (Table 24, line 5). In addition to the local effect along the ridges in the central Arctic the improvements in the EGM08 geoid are revealed in other areas such as along the high slopes of the bathymetric relief (see Figure 32c for the FM06 period).

Thus in order to assess the effect of different algorithms applied for determination of h_{sl} in the LLE and the TP methods the same scales for along-track averaging should be used to avoid corresponding bias. In order to avoid this bias when comparing the LLE and TP methods, we chose to use an averaging window of 25-km to calculate \bar{h} and h_{sl} (as in Kwok et al., 2007) for the ~~two~~ three following reasons. First, applying of the TP method using the same scales as in (Kwok et al., 2007) allow us to analyse the performance of the algorithm applied to generate the JPL product. Second, ~~U~~ sing a smaller window is found to result in reduced dependency of the freeboard on the geoid used in the retrieval process. Indeed, in this case the correction of the geoid for $\bar{h} - h_{sl}$ as well as the fact of using a recent geoid like EGM08 no longer has any impact on freeboard along the ridges in the Arctic, as opposed to what we reported above when using larger spatial averaging.

~~Another-Third~~ reason is that, as demonstrated by Kern and Spreen (2015), a more valid freeboard can be retrieved using the LLE method if the length of along-track segments considered for the selection of the lowest elevations when estimating the local sea level reference h_{sl} is equal to, or less than, the size of the smoothing window used to calculate \bar{h} values. It means that when using 100-km window for determination of h_{sl} the same (or larger) scale would be preferred to calculate \bar{h} , so that

5 fluctuations of the geoid would be not properly taken into account properly. From the other side, and as we have shown above, using a shorter length for the averaging windows when applying the LLE method results in lower freeboards that can be interpreted as underestimates due to poorer sampling of the sea level references. The TP methods, in contrast to the LLE method, is based on selection of tie-points using a physical relationship, and its performance can be assessed by comparison of the results derived from the TP and LLE approaches. As the number of ICESat measurements available within each 25-

10 km section does not exceed 147 samples, and most often even less due to data filtering, for each section the 1% of the measurements considered as the lowest h_r values and used for estimation of h_{sl} actually refer to only one sample. Kern and Spreen (2015) suggested that using such a low percentage and consequently such a low number of samples to estimate the local sea level reference may result in freeboard overestimation if sharp elevation changes are present along the track. However, this only happens if the size of the averaging window used to calculate \bar{h} is smaller than the length of the segment

15 used to estimate the local sea level reference h_{sl} , which is not the case here since we use the same 25-km averaging scales for estimation of \bar{h} and h_{sl} .

3.3 Comparison of sea ice freeboard obtained by using TP and LLE methods

3.3.1 The original algorithm used in the TP method

A comparison of sea ice freeboards calculated using the LLE and TP methods when applying the same along-track averaging scale as described in the previous section is presented in Figure 43. The obtained freeboard differences are small on average, ranging within ± 5 cm, while the presence of significant regional discrepancies should be noted. Since the maps in Figure 4 show a clear distinction between the differences over thin and thick ice for some of the ICESat periods, we estimated

20 ~~Differences between freeboards are estimated~~ separately for the first-year ice (FYI) and multi-year ice (MYI) regions over the same 25-km grid cells. Grid cells are considered as covered with FYI or MYI according to the 50% isopleth on the multi-

25 year ice fraction maps that were derived by Zygmontowska et al. (2014) from QuikSCAT scatterometer following the method described in Kwok (2004). ~~The obtained freeboard differences are small on average, ranging within ± 5 cm, while the presence of significant regional discrepancies should be noted.~~

Negative differences, which corresponds to lower freeboard being retrieved using the LLE method as compared to the TP method according to the convention used here, are found for areas covered by MYI and located north of Greenland and

30 Canadian Arctic Archipelago. The largest negative differences are observed for the period FM08 and are about -15 cm. These can be explained by the fact that in these areas of thick ice the TP method does not detect any leads for many 25-km segments, while the LLE method provides freeboard estimates because it calculates a local sea surface reference from the 1%

lowest elevations that are not necessarily representative of local sea level. One can also expect ~~basin-wide~~ difference between ~~the basin-wide~~ ~~the freeboard estimates means~~ in the area of thick MYI to be even larger when the TP method does not detect any tie-points within some ~~25 km~~ grid cells and, hence, does not provide freeboard estimates. We therefore conclude that the difference found between the results of these two methods in areas of MYI is coming from (i) the lower-

5 biased estimates due to using the measurements over refrozen leads or ice within the 25-km range for~~a poor constrain on~~ the calculation of local sea level references in the LLE method or (ii) the absence of local detection of tie-points in the TP method, mainly where the ice cover is continuous, i.e. with presence of only few or no leads.

Positive differences are obtained over large areas of FYI and thin part of MYI for most of the ICESat campaigns, with a peak in FM08. The mean differences over FYI are within 3-5 cm, while locally these can be more than 10 cm. Since the lead

10 fraction in the areas of seasonal ice is higher than over thick MYI (Willmes and Heinemann, 2016; Ivanova et al., 2016; Röhrs et al., 2012; Brohan and Kaleschke, 2014), one can expect that the observed difference in freeboard estimates is not coming from the same reasons mentioned above for MYI areas. The positive differences are most likely reflect the underestimation of freeboard retrieved by the TP method as was found by Kwok et al. (2007) from comparison with the freeboards adjacent to leads detected on satellite images and collocated with ICESat data. They showed that the freeboard

15 underestimation was on average of 1.3 to 4 cm for ON05 and FM06 periods, and explain this by the fact that samples, which are identified as tie-points, do not always represent leads or the thinnest ice. In order to explain ~~the source of why~~ freeboard differences are observed primarily over FYI areas we investigated the performance of the algorithm used in the TP method. The results are presented in the next section.

20 3.3.2 An improved algorithm for the TP method

In the TP method, as described in section 2.3, one has to find the samples that will be used as tie-points. To do so, we first establish the relationship between h_r and σ_{25} using a regression model for the measurements showing dips in reflectivity (a cubic polynomial function in Kwok et al., 2007). Figure 54 shows the relationship found between h_r and σ_{25} for the ten ICESat campaigns. We note that visually the curves corresponding to ON05 and FM06 periods are ~~We checked that our~~

25 ~~results were~~ in agreement with those reported in Kwok et al. (2007). ~~Then, a~~ According to Kwok et al. (2007), after the relationship between h_r and σ_{25} are established, the tie-points can be defined by taking the samples found to be below the regression curves. ~~The number of detected tie-points within each 25 km non overlapping segments ranges from a few tie-points (i.e. < 10) over thick MYI to several tens (i.e. > 25) over FYI (Figure 5b) and has a significant spatial variation.~~ The sea level reference for each 25-km segment is estimated by averaging the h_r values corresponding to the tie-points, weighted

30 exponentially by the distance from the regression model (Kwok et al., 2007). Hence, the contribution of tie-points with larger distances dominates when calculating the local sea level reference. The σ_{25} values are smoothed along-track and do not change remarkably over the segment, while h_r may vary significantly from sample to sample. The tie-points with lower h_r contribute more than those with larger h_r for a given σ_{25} . However, as shown by the flattening of the curves in Figure 5, ~~one~~

can note that when going toward $h_f = 0$ the quasi linear correlation existing between h_r and σ_{25} is lost as it can be seen from the flattening of the curves in Figure 6a. In FM08 this correlation becomes even inversed with an inflection point around $h_f = -5$ cm. Although when using a cubic polynomial fit of the data as in Kwok et al. (2007), reduces this flattening, the correlation does not hold towards small zero h_r for all many ICESat periods. As seen from the averaged regression lines on the Figure 5 a deviation from the linear relationship is more pronounced for the winter periods and starts in a freeboard range from -15 cm to -20 cm. However, we think that this flattening of the curves may not represent an actual and physically-based relationship existing between h_r and σ_{25} . If this is this case, some samples may be unreasonably identified as tie-points in the TP method due to enlarged area below the regression line at h_r close to zero. The effect of the flattening of the curves on the result of the regression model is illustrated on the Figure 7 for two winter periods: FM05, when the highest σ_{25} values are seen for h_r close to zero, and FM08, for which the largest discrepancy between LLE and TP results are observed. The computed regression line for the FM05 period reduces the inversed correlation between h_r and σ_{25} , but still deviates from the linear relationship (see red dashed lines in Figure 7a). In FM08 the curve is less noisy but correlation is even inversed with an inflection point around $h_r = -5$ cm, and shows an inversed correlation for FM08, or at least a flattening, as for FM05 (see red dashed lines in Figure 6a). As a consequence, the samples detected as tie-points that have a value of h_r close to zero (i.e. measurements taken over areas covered by thinner ice) may contribute more than the other, leading to an artificial increase of the reference sea level height over given segments. Note that the fact that the linear relationship between σ_{25} and h_r does not hold for thin freeboard can be explained by a lower likelihood for samples with $\Delta R > 0.3$ to represent actual leads. This is illustrated by the increase of the standard deviation of σ_{25} and the decrease in number of samples used in evaluating σ_{25} for low absolute values of h_r (Figure 7b and 7c, red). Note that this is consistent with the more pronounced flattening obtained for the winter periods, when variability of the surface roughness is larger.

Underestimation of the freeboards retrieved over thin FYI by the TP method as compared to those retrieved from the LLE method increases with the number of samples detected as tie-points and with surface roughness. This can be explained from the fact that a large number of tie-points or a high degree of roughness increase the chance that some of those tie-points would be associated with the flattening part of the curve relating σ_{25} and h_r . In general, the number of tie-points and roughness are anti-correlated and their spatial patterns match very well with the pattern of multi-year versus first-year ice as shown in Figure 6 for some selected the FM05, ON05, FM06 and FM08 ICESat periods discussed in the text, i.e. FM05, ON05, FM06 and FM08. The number of detected tie-points within each 25-km non-overlapping segments ranges from a few tie-points (i.e. < 10) over MYI to several tens (i.e. > 25) over FYI (Figure 6b) and shows a significant spatial variation. A surface roughness represented by σ_{25} (Figure 6a) typically do not exceed 10-15 cm over FYI, although it may be locally more than 20 cm as for FM05 period, when the largest roughness is observed (Figure 6a). Therefore, the differences of freeboard estimates in FM05 are primarily related to the surface roughness, and to a lesser degree to the number of detected tie-points, which is comparatively low. In contrast, the large number of detected tie points plays a key role in FM08, while surface roughness over FYI is low.

In order to reduce such bias and to ensure that the selected samples used to establish the relationship between σ_{25} and h_r are actually over leads we propose an improvement to the TP method. This improvement is based on further constraining the method of selection of samples by requesting that dips in both reflectivity and elevation need to be actually measured. Here, we select samples with $\Delta R > 0.3$ and $h_r < \overline{h_{r25}} - 0.5\sigma_{25}$, i.e. samples where h_r deviates from the 25-km running mean $\overline{h_{r25}}$ by at least half of a standard deviation. For $h_r < -15$ cm the resulting relationships between σ_{25} and h_r is very similar to the one obtained for the previous selection of samples, while for $h_r > -15$ cm both the high variability and inverse distribution are removed (Figure 76a, black). Since applying additional requirements on the selection of samples considered in the data regression reduces their number, especially for near zero h_r , we only consider as reliable the (1-cm) h_r bins for which σ_{25} is estimated from at least 15 samples. It should be noted that despite the actual lower number of samples selected with this new method the variability of σ_{25} is significantly decreased, and the relationship between h_r and σ_{25} over thin ice remains robust over the whole range of h_r and σ_{25} values. Note that we also tried to apply more stringent selection requirement on the elevation dips like e.g. $h_r < \overline{h_{r25}} - \sigma_{25}$ and $h_r < \overline{h_{r25}} - 1.5\sigma_{25}$. In this case, the resulting $\sigma_{25} = f(h_r)$ relationships (Figure 76a cyan and blue lines, respectively) are shifted downward compared to the previous one obtained when requiring $h_r < \overline{h_{r25}} - 0.5\sigma_{25}$ (Figure 76a, black line). These lines represent rather the relationships of the mean σ_{25} that Kwok et al. (2007) obtained using collocation of the satellite images with ICESat data, which we mentioned above in the section 2.3. From this analysis, using the condition $h_r < \overline{h_{r25}} - 0.5\sigma_{25}$ in this improved TP method appears to be the most appropriate [because it corrects the relationships for thin ice and, at the same time, better reproduces the TP algorithm for \$h_r < -15\$.](#)

We tested our new TP method on the whole ICESat dataset using the additional constrain $h_r < \overline{h_{r25}} - 0.5\sigma_{25}$ in the procedure of selection of samples used to form relationships between h_r and σ_{25} . The difference between the freeboards retrieved with the new-TP and LLE methods (Figure 87) is now largely reduced over FYI. Depending on the period considered, the mean difference is now varying from 1.5 cm to 3.1 cm (1.6 cm to 3.3 cm for FYI) as compared to 2.4 cm to 4.4 cm (2.5 cm to 5.3 cm for FYI) before, while the range of standard deviation of the differences remains similar. [The most remarkable improvement is observed for the FM08 period when the difference of 5-10 cm over vast areas of FYI are reduced to differences ranging within \$\pm 2\$ cm.](#) As expected, the differences remain [almost](#) unchanged for MYI since our modification of the TP method [only primarily](#) impacts freeboard estimate over thin ice areas.

3.4 Impact of snow depth in leads and lead size adjustments on sea surface reference calculation

Corrections to account for snow depth at the location of tie-points (which are supposedly leads) and for the size of leads with respect to the size of ICESat footprint (as proposed and applied by Kwok and Cunningham (2008) and Kwok et al. (2009)) is another source of contribution to the differences between the sea ice thickness products from the JPL and GSFC. The adjustment of freeboard included in the TP algorithm and related to snow depth in refrozen leads [is limited to 5 cm and](#) [is](#) about +2-3 cm [on average for all the periods considered \(Table 3, first line\).](#) [Due to the fixed limit, this correction is](#) [and](#) rather uniformly distributed over the Arctic [although we note that the lowest values are observed for thin ice in the Arctic](#)

seas, i.e. in the warmer regions with slower initial ice growth in leads (~~Table 2, first line and~~ Figure 98a for ICESat periods ON05 and FM06). The other adjustment of freeboard related to the fact that lead area does not cover the entire ICESat footprint at the locations where tie-points are detected is applied after the adjustment for snow depth previously mentioned. The magnitude of that second correction is primarily correlated to freeboard height and ranges from +3 to +7 cm on average over the Arctic depending on the period considered. Indeed, we observe that this correction is less important for the ICESat periods from and after ON05, i.e. when sea ice thickness starts to reduce significantly (Table 32, second line). Example of spatial distribution of that correction for the periods ON05 and FM06 is shown in Figure 98b. We can see that the largest corrections are observed over MYI areas. Depending on the ICESat period considered, the mean of that correction vary from 3.7 to 9.3 cm and from 2.3 to 5.6 cm over MYI and FYI areas, respectively.

The sum of these two corrections is about +7 cm on average over the Arctic and over the ICESat period, and is ranging from 5 to 10 cm depending on the particular period. Note that the mean corrections reported in Table 32 are estimated using the original TP method, and that we found very similar results when using the improved TP method that we propose in section 3.3.

In principle, the freeboards derived by the LLE method can also be corrected for snow depth in leads and lead width, but this was not done for the GSFC product. Although the LLE method selects only the lowest elevations to determine the local sea level, these samples may be contaminated by snow accumulated in leads or by the neighbouring sea ice surface within the laser altimeter footprint. As shown above from the comparison of freeboards derived by the LLE and TP methods, the LLE method has a weakness mostly over the thickest ice due to lack of leads. The GSFC product, in addition, was derived using longer averaging windows that, in general, increases a likelihood for the lowest elevations to represent a true height of sea level. Therefore, we think that when using the LLE method the empirical functions proposed by Kwok and Cunningham (2008) and Kwok et al. (2009) should be modified depending on the averaging scale applied.

3.5 Summary

Mean sea ice freeboard calculated over the whole arctic basin using different methods is shown in the Figure 109. The figure includes freeboard estimates from the GSFC product (black), estimates we calculated using the same LLE method (but with finer resolution along-track averaging) (red), those we calculated using the original TP method as described in Kwok et al. (2007) (green), and those calculated using the improved TP method we propose in this study, with (blue) or without (green dashed) adjustments for snow depth and leads width. Note that these results correspond to the area considered in the JPL product, i.e. to the Arctic Ocean without surrounding Arctic seas such as Greenland Sea, Barents Sea, Kara Sea and Baffin Bay. Correspondingly, the difference of 0.42 m between sea ice thickness in the JPL and GSFC products found by Lindsay and Schweiger (2015) was derived using randomly selected samples over Arctic Basin. Although excluding of the above-mentioned seas does not significantly impacts the difference between the results, the freeboard means are changed within the range of ± 2 cm depending of the ICESat period. The average freeboards obtained from the different methods and corresponding to the results shown on Figure 10 are recaped in Table 4.

~~for consistency in the comparison we computed mean values of freeboard by only considering grid cells where estimates were available for all these methods.~~

Because of the use of different averaging scales to calculate sea level references, the sea ice freeboards we estimated using the LLE method with a 25-km averaging window are lower by $\sim 4\text{--}3$ cm on average as compared to those of the GSFC product for all ICESat periods (Figure 10, red and black lines). As we discussed in section 3.2, this can be explained by the increased likelihood for the lowest elevations to correspond to actual sea level height when using a larger window (e.g. 100 km as applied to produce the GSFC dataset).

The S sea ice freeboards we estimated using the original TP method are lower by ~ 3 cm on average as compared to those we obtained ~~using~~ with the LLE method with identical 25-km along-track averaging scales (Figure 10, green and red lines).

~~Although the TP method has the advantage of avoiding bias due to the use of samples over sea ice to calculate the local sea level reference, and prevent from local underestimation of the freeboard in thick ice areas, it may underestimate freeboard estimates over the large areas covered by thin ice. As already shown by Kwok et al. (2007), this study shows that the tie-points for determination of sea level references selected by the TP method not always represent leads or the thinnest ice. Also, the freeboards obtained by the TP method are lower over MYI for the first three ICESat periods (ON03, FM04 and ON04) where medium thick freeboard of about 30 cm is observed.~~

~~The lower average freeboard retrieved by the TP method can be explained by using the large number of tie points found in thin ice areas. Although the tie-points with lower elevations have greater weight in most cases, the resulting sea surface reference is biased positive, hence leading to lower freeboard estimates (see Figure 5 and 7). Therefore, we suppose that obtained difference between the TP and LLE results reflects the underestimation of the freeboard derived by the TP method.~~

The lower freeboard obtained by the TP method is mostly observed over FYI and part of MYI in the central Arctic especially in the first three ICESat periods (ON03, FM04 and ON04) (Figures 4 and 10). This was already seen in Kwok et al. (2007) as the freeboard retrieved by the TP method for ON05 and FM06 periods were, in average, lower by 1.3 to 4 cm as compared to freeboards neighboring to leads detected in parallel on satellite images and collocated with ICESat data.

However, the difference between freeboards retrieved by the LLE and TP methods is reduced by more than 30% for the whole Arctic and by 40% for FYI when applying our suggested improvements of the algorithm used in the TP method. At the same time, over thick part of MYI the LLE method tends to give lower freeboards. Although it is not reflected in the mean values on the Figure 10, it can be seen on the maps of the differences (Figure 4 and 8), especially for the FM08 period. This is consistent with our expectations that over thick part of continuous MYI with fewer leads, the use of a relationship between the freeboard and surface roughness for identification of tie-points, as done in the TP method, gives more reliable freeboard estimates.

The adjustments for snow depth in leads and leads width were applied only when producing the JPL dataset, and we estimated that their combined effect increases freeboard by about 7 cm. Although the samples used for determination of sea level in the LLE method can also be affected by snow accumulation and contaminated by the neighbouring sea ice surface, these corrections were not applied to freeboards in the GSFC product. Therefore, S since the ratio between mean ~~sea ice~~ total

freeboard and thickness reported in Kwok et al. (2009) is about 6, the application of these two adjustments could be in principle sufficient to explain the difference of 0.42 m on average found between the JPL and GSFC sea ice thickness products (Lindsay and Schweiger, 2015). However our results show that after applying both corrections the freeboards retrieved using the original and improved TP method are, on average, similar and higher by ~1 cm respectively, as compared to the GSFC product. This means that, on average, it is sufficient to account for these corrections by using a larger averaging window and LLE method or that both products underestimate freeboard.

As sea ice freeboard data from JPL are not available we cannot check their consistency with those estimated for this study using the same method, or provide a comparison with the estimates provided in the GSFC product. According to our findings the freeboards of the GSFC product and those that were most likely calculated at JPL are close ~~o~~ⁱn average, meaning that the difference between the JPL and GSFC averaged sea ice thicknesses are probably coming from the difference in the choice of parameter values used in the freeboard-to-thickness conversion.

Conclusions

In this paper we reproduced two methods already used in other studies to retrieve sea ice freeboard using ICESat data, called the lowest level elevation (LLE) and tie-points (TP) methods. The main difference between retrieval algorithms used in these two methods reside in applying different approaches to determine the local sea level references – a key step in the process of estimating sea ice freeboard. Two available products of the Arctic sea ice thickness, GSFC and JPL, were derived respectively from freeboards retrieved with these two approaches, but were found to differ significantly, i.e. by 0.42 m (Lindsay and Schweiger, 2015). In this study we analyzed the possible reasons for freeboard discrepancies when using the LLE and TP methods as well as their contribution to the observed thickness difference.

We first reproduced the freeboard estimates of the GSFC product by using the algorithm of the LLE method. We estimated the contribution of using different along-track averaging scales in the TP and LLE methods (as it is the case between Kwok et al. (2007) and Yi and Zwally (2009), respectively) on the freeboard estimation and how it could possibly explain the sea ice thickness differences found between the JPL and GSFC products. The along-track averaging scales used by Yi and Zwally (2009) are larger than those used by Kwok et al. (2007), resulting in higher freeboard estimates by ~43 cm ~~o~~ⁱn average. We also estimated the effect of the geoid adjustment applied by Yi and Zwally (2009) for the residuals between the geoid heights and the sea level determined from ICESat data. We found only local ~~In this analysis,~~ noticeable freeboard differences are observed only locally, i.e. along the high slopes of the bathymetric relief, and only when using large along-track averaging scale.

In order to analyse the effect of using different approaches to estimate local sea level references, the same 25-km along-track averaging was applied for both the LLE and TP methods. ~~Our results~~We showed that locally and over thick and continuous MYI cover areas ~~with very few leads~~ the LLE method gives a lower sea ice freeboard by up to 15 cm when compared to the TP method. Over FYI, in contrast, the LLE method gives sea ice freeboards that are higher by 3-5 cm ~~i~~^on average compared to the TP method. This ~~can be~~^{is} explained by the ~~large number of~~fact that ICESat samples selected for calculating local sea

level do not always represent the lowest elevations and by their inadequate weighting in these calculations when applying the TP method. ~~In this study, w~~We proposed an improvement in the algorithm of the TP method that results in a much better agreement over FYI with the LLE method, i.e. with differences reduced to less than 2 cm on average. Since it is based on a physical relationship and seems adequate to give reasonable results over both MY and FY areas, we therefore

5 ~~Therefore, we would~~ recommend using the TP method with our improved algorithm over the LLE method to calculate local sea surface references ~~as it is based on a physical relationship and seems adequate to give reasonable results over both MY and FY areas.~~

The freeboard corrections that have been applied in ~~the last release of~~ the JPL product to account for snow depth in leads and for leads ~~width~~ with respect to the size of the ICESat altimeter footprint (Kwok and Cunningham, 2008; Kwok et al., 2009)
10 are significantly impacting the freeboard values estimated using the TP method, accounting for an increase of about 7cm ~~on~~ average.

Overall, Wwe showed that the different along-track averaging scales and approaches to calculate sea surface references, from one side, and the freeboard adjustments as applied in the TP method used to produce the JPL dataset, from the other side, are roughly compensating each other with respect to freeboard estimation. Indeed, we obtain similar freeboard
15 estimates while using the TP and LLE methods with the set of parameters used by Kwok et al. (2009) and Yi and Zwally 2009, respectively. We therefore suspect that the differences found in the JPL and GSFC sea ice thickness products are not intrinsically due to the difference in the freeboard retrieval methods, but may be attributed to the use of differences in the freeboard-to-thickness conversion.

In conclusion, we show that using different methods for sea ice freeboard retrieval from ICESat data leads to persistent
20 differences between freeboard estimates over large areas in the Arctic. In particular, significant freeboard biases are found when using different algorithms and averaging scales for determination of the sea surface height, to which the freeboard is referenced. These biases depend on sea ice characteristics, such as lead fraction and surface roughness, and therefore vary in space and time. In addition, the freeboard adjustments accounting for snow depth in leads and lead width in the JPL product significantly affect freeboard values. However, it is difficult to assess the validity of the empirical relationships proposed for
25 calculation of these corrections by Kwok and Cunningham (2008) and Kwok et al. (2009) and whether they should be applied to the freeboards retrieved by the LLE method using larger averaging scale as it was done for the GSFC product. We demonstrate and quantify the sources of uncertainties in the freeboard retrieval process, and propose the improvement of the TP method that can be used for further studies related to the freeboard retrieval from satellite laser altimetry. Although ICESat-1 is currently not in orbit anymore these findings can be used in the future studies for analysis of the data from the
30 follow up ICESat-2 satellite planned for launch in 2018.

Acknowledgments. This work was funded and supported by the Nansen Environmental and Remote Sensing Center and by the ESA Climate Change Initiative, Sea Ice project (SICCI).

References

- 5 [Bröhan, D., Kaleschke, L.: A Nine-Year Climatology of Arctic Sea Ice Lead Orientation and Frequency from AMSR-E. Remote Sensing 6 \(2\), 1451-1475, doi:10.3390/rs6021451, 2014.](#)
- [Cavalieri, D. J., T. Markus, and J. C. Comiso. AMSR-E/Aqua Daily L3 25 km Brightness Temperature & Sea Ice Concentration Polar Grids, Version 3. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: \[http://dx.doi.org/10.5067/AMSR-E/AE_SI25.003\]\(http://dx.doi.org/10.5067/AMSR-E/AE_SI25.003\), 2014.](#)
- 10 Farrell S. L., Laxon S. W., McAdoo D. C., Yi D and Zwally H. J.: Five years of Arctic sea ice freeboard measurements from the Ice, Cloud and land Elevation Satellite. *J. Geophys. Res.*, 114(C4):C04008. doi:10.1029/2008JC005074, 2009.
- [Ivanova, N., Rampal, P., and Bouillon, S.: Error assessment of satellite-derived lead fraction in the Arctic, The Cryosphere, 10, 585-595, doi:10.5194/tc-10-585-2016, 2016.](#)
- Kern S. and Spreen G.: Uncertainties in Antarctic sea-ice thickness retrieval from ICESat, *Ann. Glaciol.*, 56(69), 107-119, doi:10.3189/2015AoG69A736, 2015.
- 15 Kurtz N. T. and Markus T.: Satellite observations of Antarctic sea ice thickness and volume, *J. Geophys. Res.*, 117, C08025, doi:10.1029/2012JC008141, 2012.
- Kurtz N. T., Markus T., Cavalieri D. J., Sparling L. C., Krabill W. B., Gasiewski A. J. and Sonntag J. G.: Estimation of sea ice thickness distributions through the combination of snow depth and satellite laser altimetry data. *J. Geophys. Res. Ocean*, 114, C10007, doi:10.1029/2009JC005292, 2009.
- 20 Kurtz N. T., Markus T., Farrell S. L., Worthen D. L. and Boisvert L. N.: Observations of recent Arctic sea ice volume loss and its impact on ocean-atmosphere energy exchange and ice production, *J. Geophys. Res.*, 116, C04015, doi:10.1029/2010JC006235, 2011.
- Kwok R.: Annual cycles of multiyear sea ice coverage of the Arctic Ocean: 1999–2003, *J. Geophys. Res.*, 109, C11004, doi:10.1029/2003JC002238, 2004.
- 25 Kwok R. and Cunningham G. F.: ICESat over Arctic sea ice: estimation of snow depth and ice thickness. *J. Geophys. Res.*, 113(C8), C08010, doi: 10.1029/2008JC004753, 2008.
- Kwok R, Zwally H. J. and Yi D.: ICESat observations of Arctic sea ice: a first look. *Geophys. Res. Lett.*, 31(16), L16401 (doi:10.1029/2004GL020309), 2004.
- Kwok R., Cunningham G. F., Zwally H. J. and Yi D.: ICESat over Arctic sea ice: Interpretation of altimetric and reflectivity profiles, *J. Geophys. Res.*, 111, C06006, doi:10.1029/2005JC003175, 2006.
- 30 Kwok R., Cunningham G. F., Zwally H. J. and Yi D.: Ice, Cloud, and land Elevation Satellite (ICESat) over Arctic sea ice: Retrieval of freeboard, *J. Geophys. Res.*, 112, C12013, 2007.

- Kwok R., Cunningham G. F., Wensnahan M., Rigor I., Zwally H. J. and Yi D.: Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008, *J. Geophys. Res.*, 114, C07005, doi:10.1029/2009JC005312, 2009.
- Laxon S., Peacock N. and Smith D.: High interannual variability of sea ice thickness in the Arctic region, *Nature*, 425(6961), 947-950, 2003.
- 5 Laxon S. W., Giles K. A., Ridout A. L., Wingham D. J., Willatt R., Cullen R., Kwok R., Schweiger A., Zhang J., Haas C., Hendricks S., Krishfield R., Kurtz N., Farrell S. and Davidson M.: CryoSat-2 estimates of Arctic sea ice thickness and volume, *Geophys. Res. Lett.*, 40, 732-737, doi:10.1002/grl.50193, 2013.
- Lindsay R. and Schweiger A.: Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations, *The Cryosphere*, 9(1), 269-283, doi:10.5194/tc-9-269-2015, 2015.
- 10 Markus T., Massom R., Worby A., Lytle V., Kurtz N. and Maksym T.: Freeboard, snow depth and sea-ice roughness in East Antarctica from in situ and multiple satellite data, *Ann. Glaciol.*, 52(57), 242-248, 2011.
- Pavlis N. K. , Holmes S. A., Kenyon S. C. and Factor J. K.: The development and evaluation of the Earth Gravitational Model 2008 (EGM2008), *J. Geophys. Res.*, 117, B04406, doi:10.1029/2011JB008916, 2012.
- Ricker R., Hendricks S., Helm V., Skourup H. and Davidson M.: Sensitivity of Cryosat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation, *The Cryosphere*, 8, 1607–1622, doi:10.5194/tc-8-1607-2014, 2014.
- 15 [Röhrs, J., Kaleschke, L., Bröhan, D., and Siligam, P. K.: Corrigendum to “An algorithm to detect sea ice leads by using AMSR-E passive microwave imagery” published in *The Cryosphere*, 6,343–352, *The Cryosphere*, 6,365–365, 2012.](#)
- [Spreen G, Kern S, Stammer D, Forsberg R and Haarpaintner J: Satellite-based estimates of sea-ice volume flux through Fram Strait, *Ann. Glaciol.*, 44, 321–328, 2006, doi: 10.3189/ 172756406781811385.](#)
- 20 Stocker T. F., Dahe Q. and Plattner G. K.: Climate Change 2013: The Physical Science Basis, Intergovernmental Panel on Climate Change 2013, 2013.
- Tilling R. L., Ridout A., Shepherd A. and Wingham, D. J.: Increased Arctic sea ice volume after anomalously low melting in 2013, *Nature Geoscience*, 8, 643-646, 2015.
- Vaughan D., Comiso J., Allison I., Carrasco J., Kaser G., Kwok R., Mote P., Murray T., Paul F., Ren J., Rignot E., Solomina O., Steffen K., and Zhang T.: Observations: Cryosphere, book section 4, 317–382, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.012, www.climatechange2013.org, 2013.
- 25 [Willmes, S. and Heinemann, G.: Sea-ice wintertime lead frequencies and regional characteristics in the Arctic, 2003-2015, *Remote Sensing, Special Issue: Sea Ice Remote Sensing and Analysis*, 8\(1\), 4, doi: 10.3390/rs8010004, 2016.](#)
- 30 [Xie, H., Tekeli A., Ackley S., Yi D. and Zwally J.: Sea ice thickness estimations from ICESat Altimetry over the Bellingshausen and Amundsen Seas, 2003-2009, *J. Geophys. Res.*, doi:10.1002/jgrc.20179, 2013](#)
- Yi D. and Zwally H. J.: Arctic Sea Ice Freeboard and Thickness, National Snow and Ice Data Center, Boulder, Colorado, USA, 2009, 2014-04-15.

Yi D., Zwally H. J. and Robbins J. W.: ICESat observations of seasonal and interannual variations of sea-ice freeboard and estimated thickness in the Weddell Sea, Antarctica (2003 – 2009), [Annals of Glaciology](#), 52, 43-51, 2011.

[Zwally, H. J., Schutz, B., Abdalati, W., Abshire, J., Bentley C., Brenner A., Bufton J., Dezio J., Hancock D., Harding D., Herring T., Minster, B., Quinn K., Palm S., Spinhirne J., Thomas R., 2002. ICESat's Laser Measurements of Polar Ice, Atmosphere, Ocean, and Land. Journal of Geodynamics 34\(3-4\): 405-445.](#)

5

Zwally H. J., Yi D., Kwok R. and Zhao Y.: ICESat measurements of sea ice freeboard and estimates of sea ice thickness in the Weddell Sea, J. Geophys. Res., 113, C02S15, doi:10.1029/2007JC004284, 2008.

Zygmuntowska M., Rampal P., Ivanova N. and Smedsrud L. H.: Uncertainties in Arctic sea ice thickness and volume: new estimates and implications for trends, The Cryosphere, 8, 705-720, doi:10.5194/tc-8-705-2014, 2014.

Table 1. Methods and settings used for freeboard retrieval in the GSFC and JPL products.

<u>Institute</u>	<u>Method to estimate h_{sl}</u>	<u>Segment length for \bar{h}, km</u>	<u>Segment length for h_{sl}, km</u>
<u>GSFC</u>	<u>LLE (lowest 1%)</u>	<u>50</u>	<u>100</u>
<u>JPL</u>	<u>TP (using $\sigma_{25} = f(h_r)$)</u>	<u>25</u>	<u>25</u>

Table 2+. Mean \pm std of the differences between freeboards estimated using different averaging scales and geoids (cm). The compared methods are indicated by the used geoid (ArcGP or EGM08), geoid adjustment is indicated by dif, scale of along-track averaging is indicated by 50 (i.e. applying of 50-km and 100-km windows for estimating of \bar{h} and h_{sl}) and 25 (i.e. applying of 25-km windows for estimating of \bar{h} and h_{sl}).

Methods compared	ON03	FM04	ON04	FM05	ON05	FM06	ON06	MA07	ON07	FM08
ArcGP/dif/50 – GSFC	0.3 \pm 6.5	0.4 \pm 5.6	0.6 \pm 6.7	1.2 \pm 7.3	- 1.5 \pm 6.6	0.1 \pm 5.7	- 1.7 \pm 6.5	- 0.3 \pm 5.9	- 2.0 \pm 7.2	- 0.4 \pm 5.7
ArcGP/dif/50 (gain 80) – GSFC	4.6 \pm 12. 4	1.3 \pm 5.5	4.6 \pm 9.2	3.1 \pm 8.6	1.9 \pm 6.3	1.7 \pm 5.5	1.4 \pm 5.5	1.5 \pm 5.5	2.8 \pm 6.0	1.2 \pm 5.2
EGM08/50 – EGM08/25	5.2 \pm 8.8	3.6 \pm 7.3	5.7 \pm 8.4	6.4 \pm 9.1	3.7 \pm 7.8	3.8 \pm 7.5	3.3 \pm 6.6	3.8 \pm 7.7	4.1 \pm 8.4	3.8 \pm 8.6
EGM08/dif/50 – EGM08 / 50	0.1 \pm 5.2	0.7 \pm 4.2	0.0 \pm 5.2	0.2 \pm 7.0	0.3 \pm 4.8	0.7 \pm 3.9	0.3 \pm 4.6	0.5 \pm 4.4	0.0 \pm 4.7	0.5 \pm 3.8
EGM08/50 – ArcGP /50	0.0 \pm 3.7	0.2 \pm 2.7	0.0 \pm 3.2	0.2 \pm 2.7	0.1 \pm 3.0	0.2 \pm 2.8	0.1 \pm 3.1	0.2 \pm 2.7	0.0 \pm 5.9	0.2 \pm 2.7

Table 32. Mean±std of the Adjustments of sea ice freeboard retrieved by the TP method to account for snow depth in refrozen leads and for leads size with respect to the size of the ICESat footprint (cm). The adjustments are estimated following the methods described in Kwok and Cunningham (2008) and Kwok et al., (2009).

Correction	ON03	FM04	ON04	FM05	ON05	FM06	ON06	MA07	ON07	FM08
Snow depth	2.5 ± _{±1.}	3.0 ± _{±1.}	2.8 ± _{±1.}	3.0 ± _{±1.}	2.7 ± _{±1.}	2.7 ± _{±1.}	2.7 ± _{±1.}	2.9 ± _{±1.}	2.3 ± _{±1.}	2.6 ± _{±1.}
	0	0	0	0	1	1	2	2	3	1
Leads width	4.5 ± _{±2.}	3.1 ± _{±1.}	6.1 ± _{±3.}	6.9 ± _{±3.}	4.1 ± _{±2.}	4.1 ± _{±2.}	3.3 ± _{±1.}	3.6 ± _{±1.}	3.2 ± _{±1.}	2.6 ± _{±1.}
	3	7	1	3	4	2	9	9	8	6

Table 4. Mean freeboard as derived from the GSFC product and estimated in this study using different methods, along-track averaging scales, and with or without applied corrections for snow depth in leads and lead width.

<u>Area</u>	<u>GSFC</u>	<u>LLE,</u> <u>100 km</u>	<u>LLE,</u> <u>25 km</u>	<u>TP</u>	<u>TP +</u> <u>corrections</u>	<u>TP modified</u>	<u>TP modified</u> <u>+</u> <u>corrections</u>
<u>Overall</u>	<u>32.7</u>	<u>33.0</u>	<u>29.5</u>	<u>26.4</u>	<u>33.4</u>	<u>27.3</u>	<u>34.5</u>
<u>FYI</u>	<u>24.2</u>	<u>24.4</u>	<u>20.7</u>	<u>17.6</u>	<u>23.5</u>	<u>18.9</u>	<u>24.9</u>
<u>MYI</u>	<u>43.4</u>	<u>44.3</u>	<u>40.3</u>	<u>38.5</u>	<u>46.9</u>	<u>38.8</u>	<u>47.3</u>

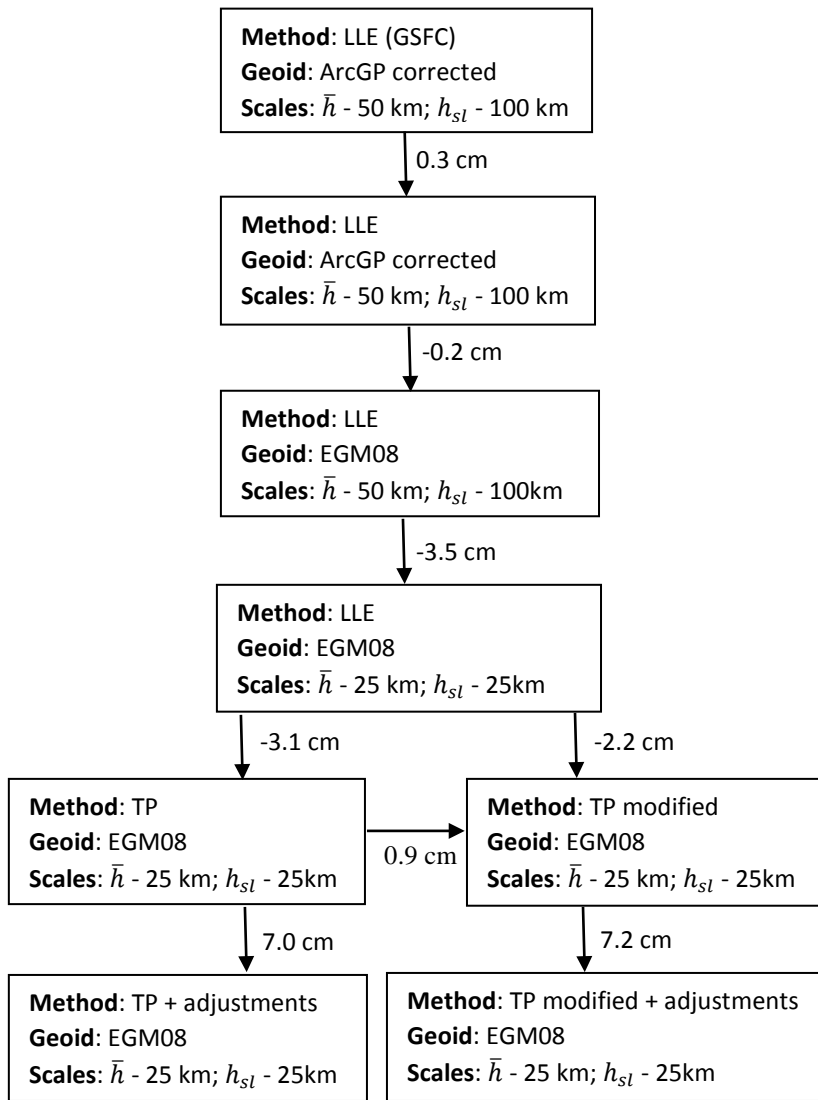


Figure 1. Flow chart of the methods, settings and geoids used for freeboard retrieval in our implementations of the TP and LLE methods. Values next to the arrows are the mean differences between freeboard estimates as derived from the Table 4.

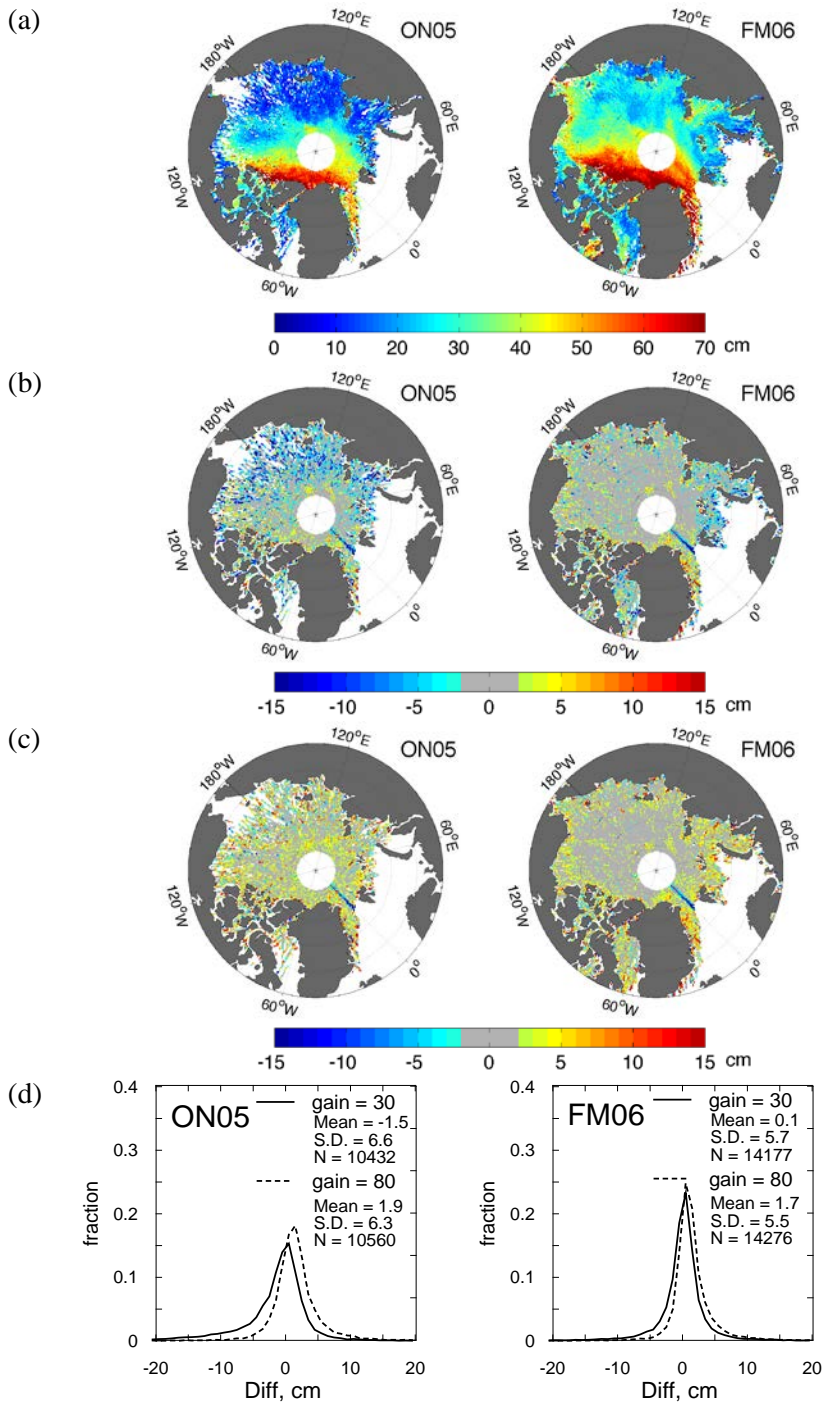
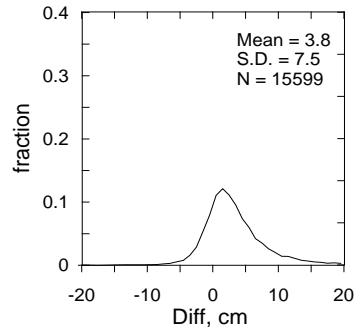
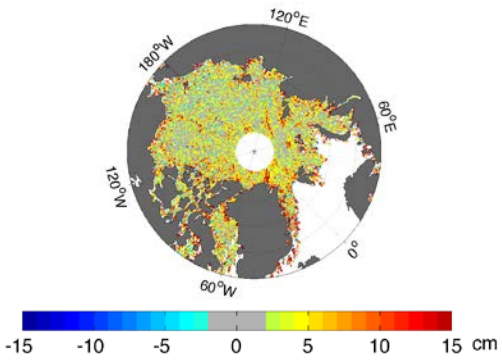
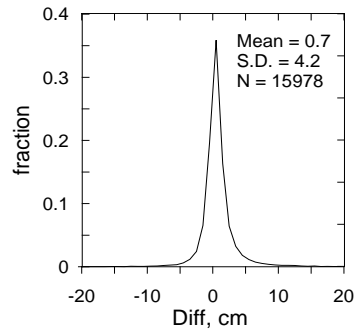
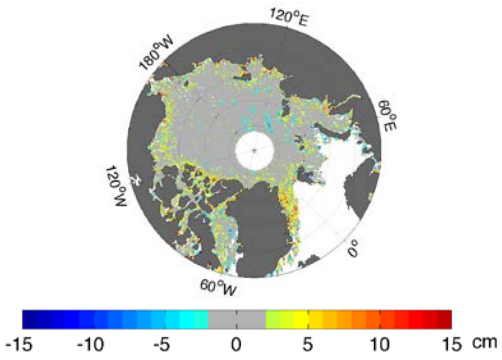


Figure 24. (a) Freeboard retrieved by the LLE method using along-track averaging scales as applied by Yi and Zwally (2009) to derive the GSFC product and (b-d) its differences from the GSFC freeboards (freeboards from this study minus GSFC freeboards) for ON05 and FM06 periods gridded into 25-km bins (cm). The freeboards estimated in this study are obtained using ICESat data with receiver gain of smaller than 30 (b) and 80 (c). (d) Distribution of the differences between freeboards. [An artefact line of negative differences along the 0 longitude is due to an unexplained positive anomaly in the GSFC freeboard estimates.](#)

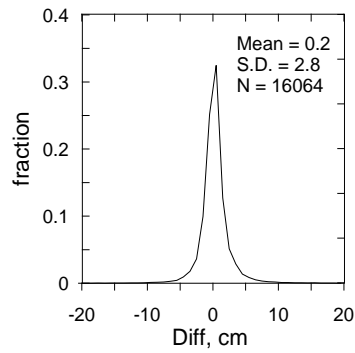
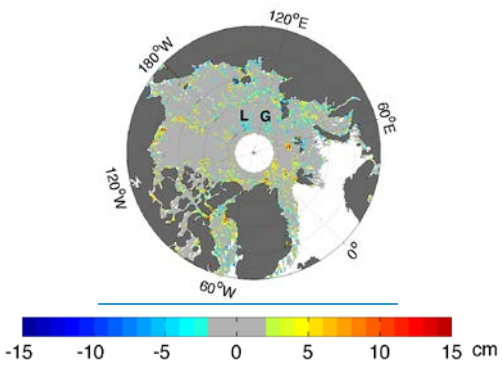
(a)



(b)



(c)



(d)

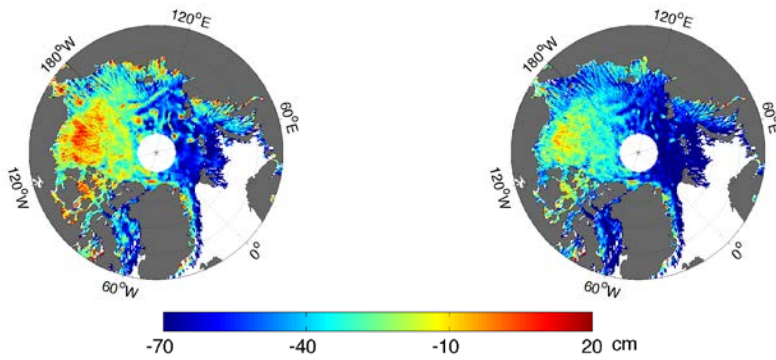


Figure 32. Effect of using different geoids and scales for along-track averaging on freeboard estimates (cm) when applying the LLE method for FM06 period (25-km grids). Differences between freeboard estimates show the effects of (a) applying longer and shorter along-track averaging scales (longer minus shorter) when using EGM08 geoid as well as (b) adjustment of ArcGP geoid for $\bar{h} - h_{sl}$ values (adjusted minus unadjusted) and (c) using different geoids (EGM08 minus ArcGP) in case of applying longer along-track averaging scales. (d) Adjustment for $\bar{h} - h_{sl}$ to correct ArcGP (left) and EGM08 (right) geoids when using coarser resolutions (cm). G and L in (c) point out on the location of the differences along the Gakkel and Lomonosov ridges.

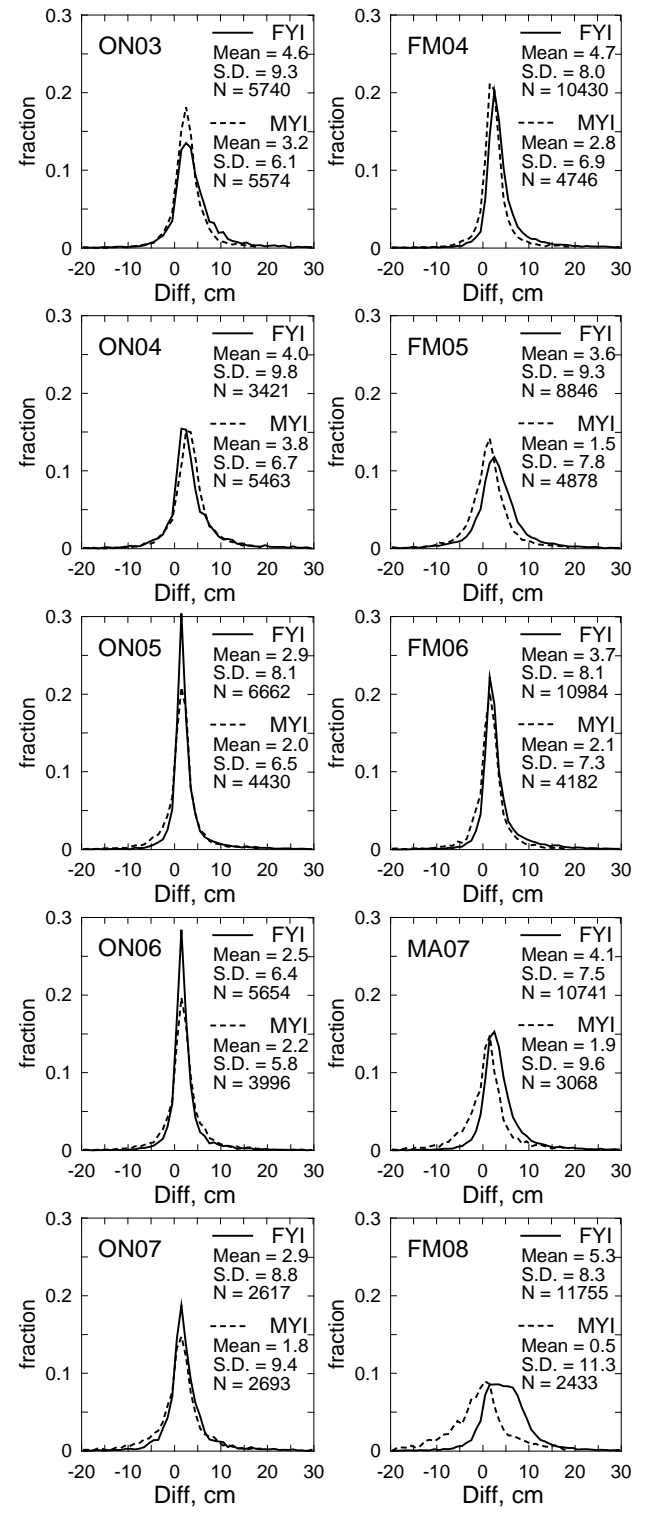
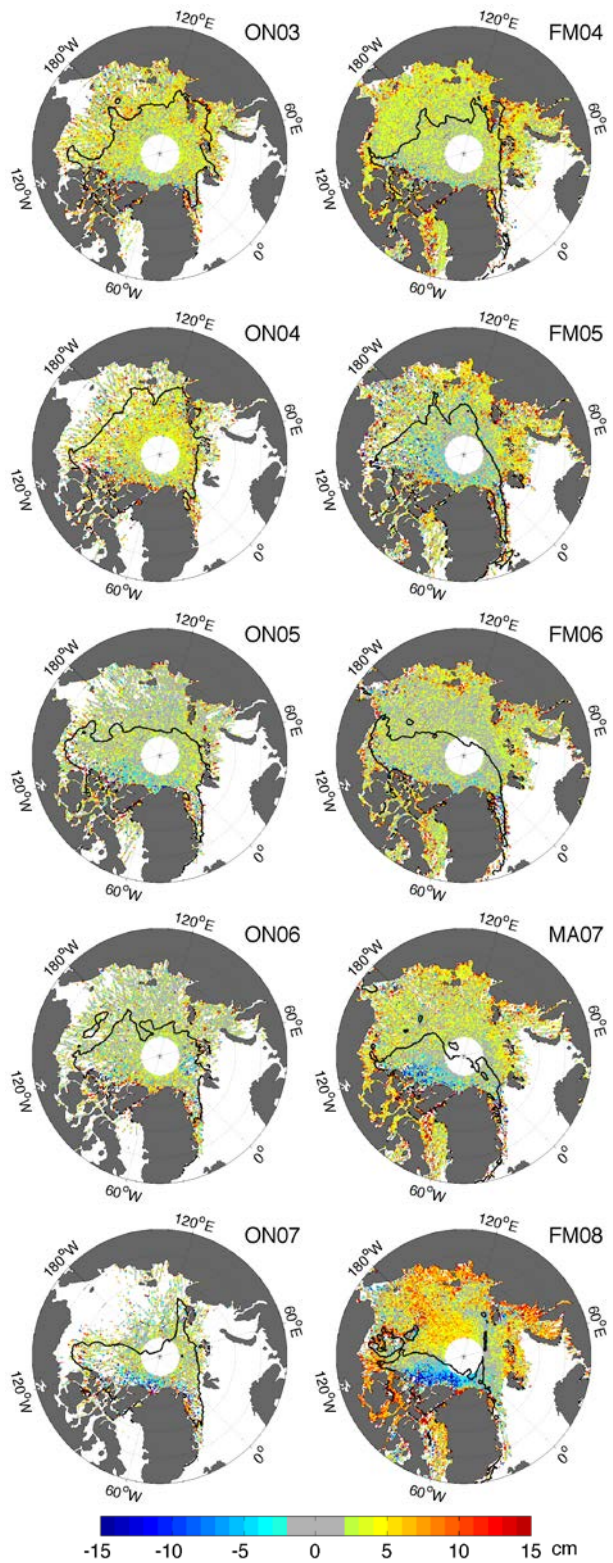
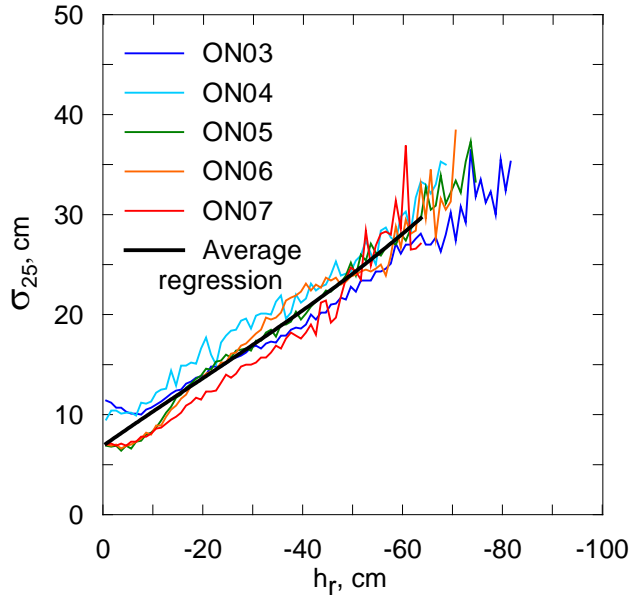


Figure 43. Maps (left) and distributions (right) of the differences between sea ice freeboard (25-km grids) estimated using LLE and TP methods (LLE minus TP) for ten ICESat periods (cm). Thick black line on the maps delineates the average 50% isopleth of multi-year ice fraction.

5

(a)



(b)

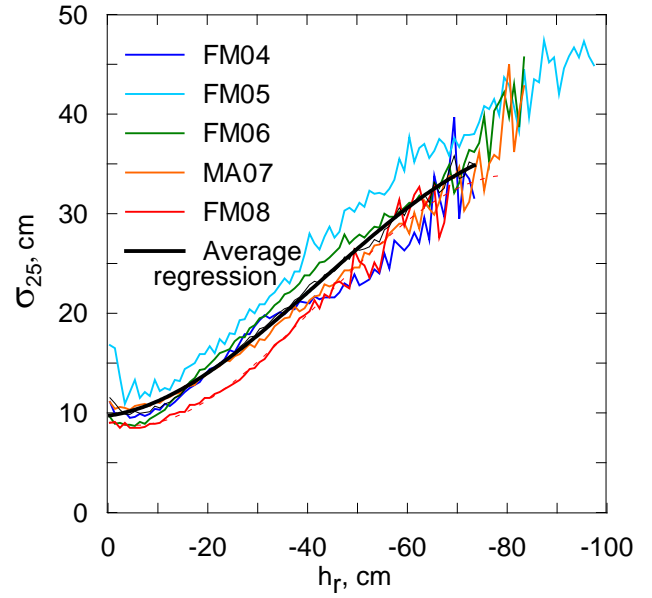
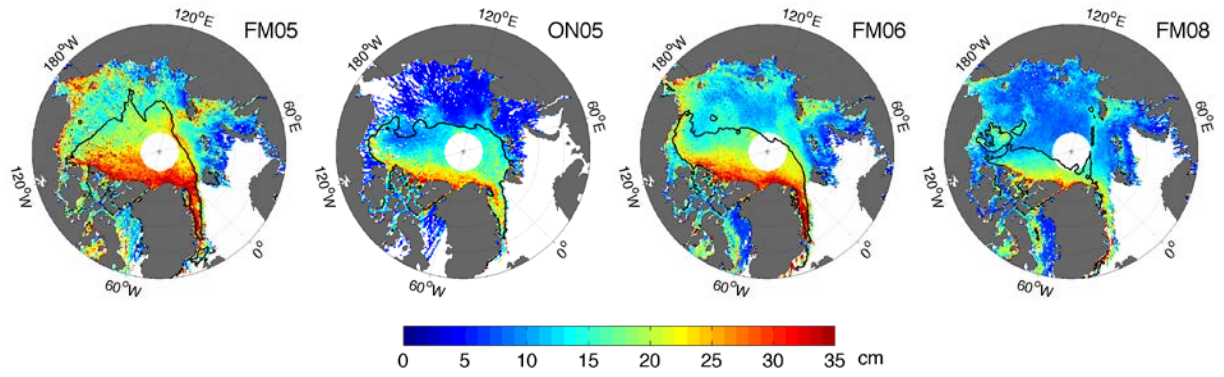


Figure 54. The relationships between h_r and σ_{25} , for samples where dips in reflectivity is measured for fall (a) and winter (b) ICESat periods, following the method described in Kwok et al. (2007). The h_r axis is discretized in bins of 1-cm. Note that following Zwally et al. (2008) and Yi and Zwally (2009) we define h_r as $h_r = h - \bar{h}$ and form the relationship for negative h_r values, while in Kwok et al. (2007) $h_r = \bar{h} - h$ and positive h_r values are considered.

(a)



(b)

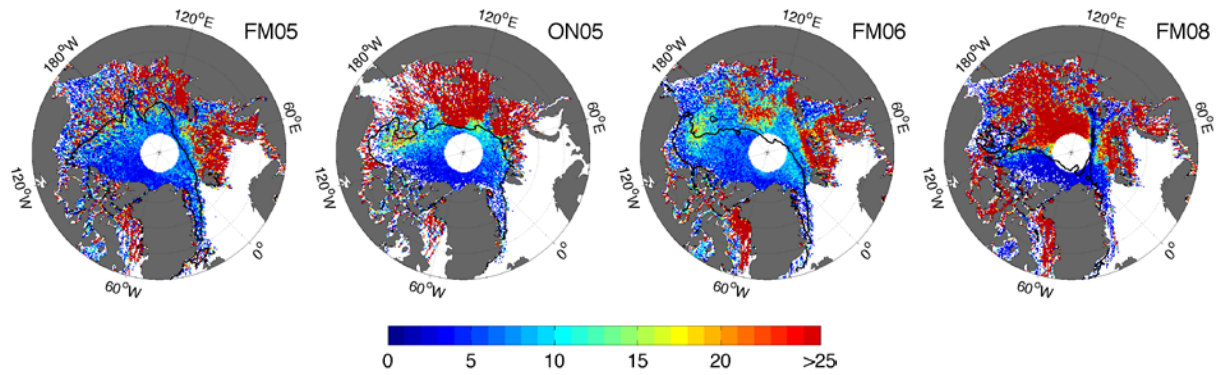


Figure 65. Standard deviation of detrended elevations h_r (cm) (a) and number of tiepoints within 25-km non-overlapping segments detected by the TP method (b) for FM05, ON05, FM06 and FM08 periods.

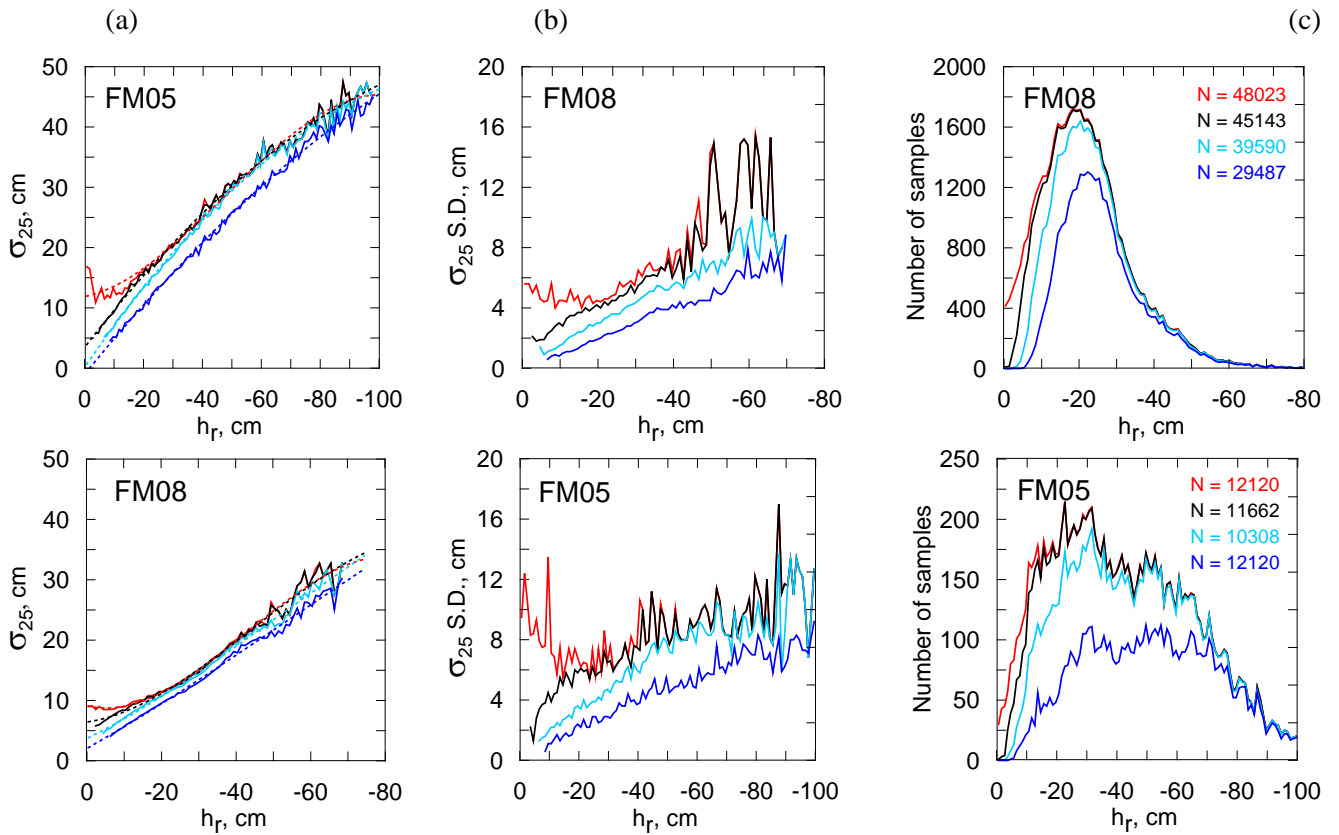


Figure 76. The relationship between h_r and σ_{25} (a), distributions of its standard deviations (b), and number of samples in each h_r bin (c) for FM05 and FM08 ICESat periods. Red lines are constructed from the selection of samples for which dips in reflectivity are measured, as described in (Kwok et al., 2007). Black, cyan and blue lines are constructed from the new selection of samples we propose in this study, based on requesting the presence of dips in both reflectivity and elevation measurements. [Dashed lines correspond to the result of the regression model \(cubic polynomial fits\) applied to the data](#) (See text for details).

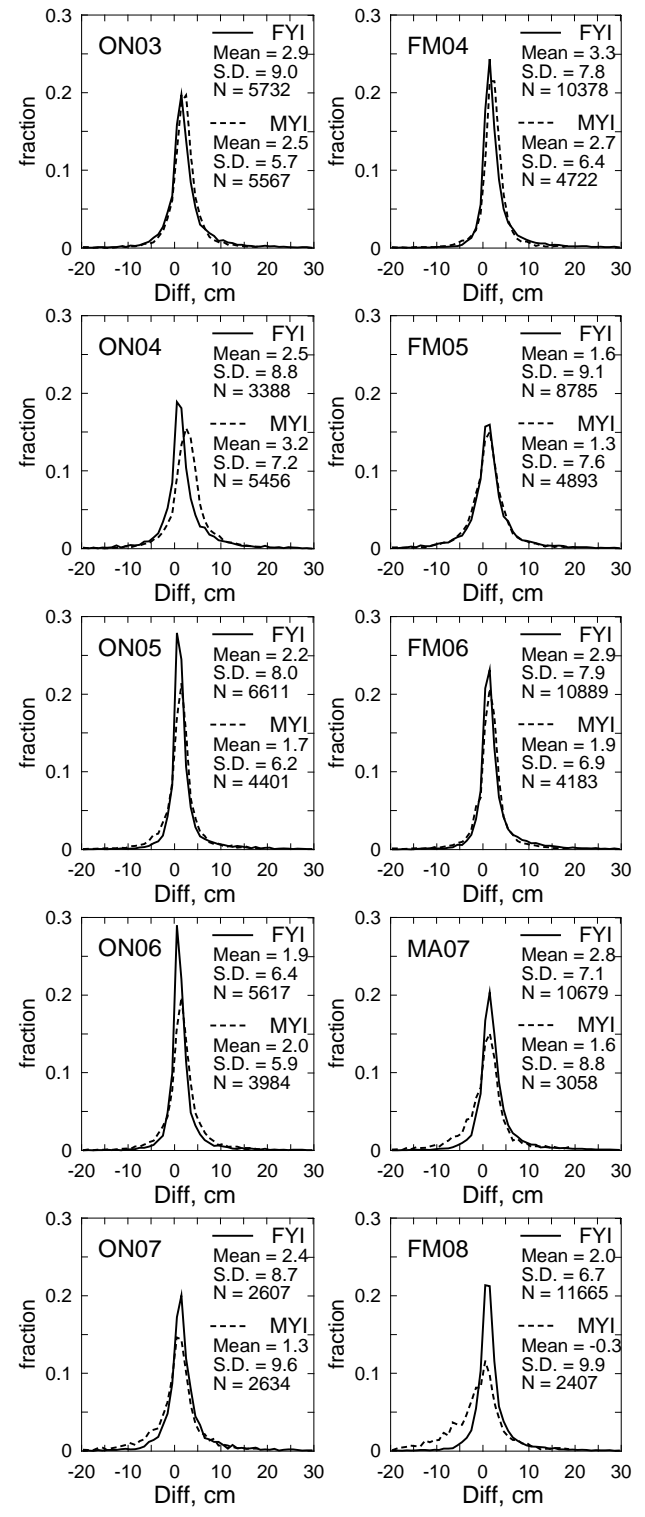
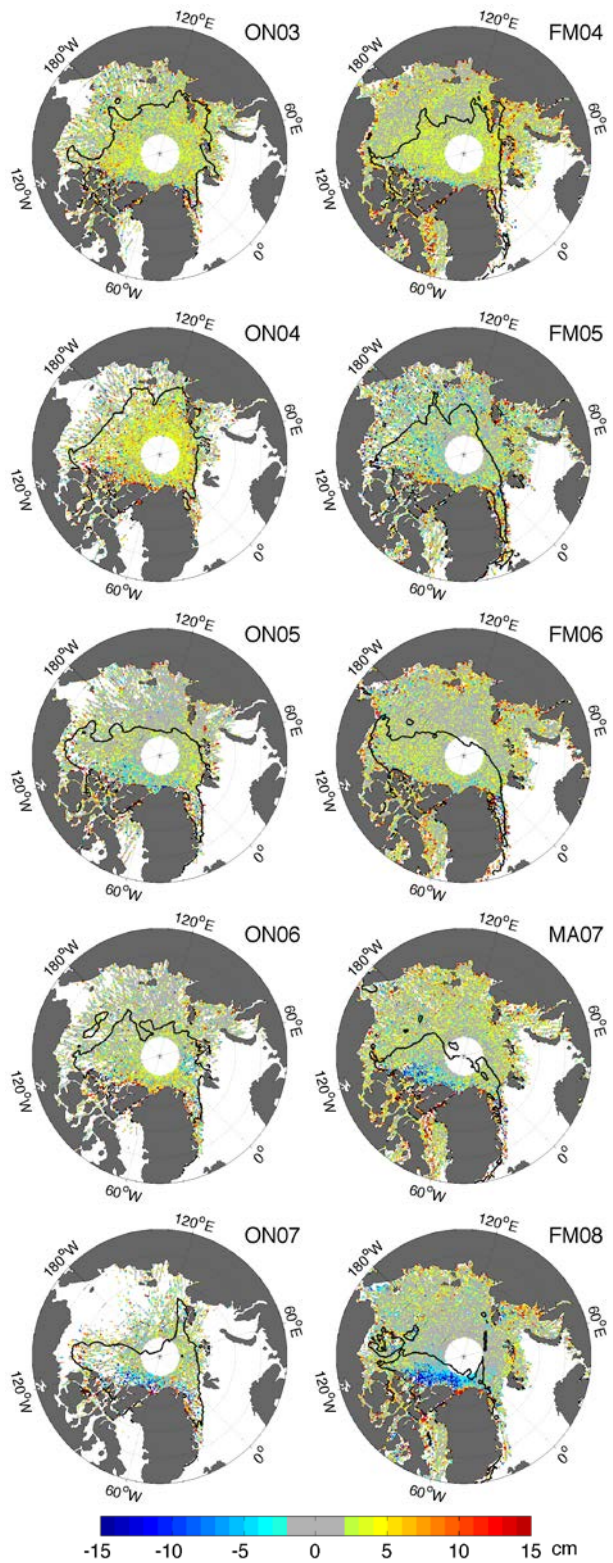


Figure 8.7. Maps (left) and distributions (right) of the differences between sea ice freeboard (25-km grids) estimated using LLE and TP methods (LLE minus TP) for ten ICESat periods (cm). The TP method used here includes the improvements in the freeboard retrieval algorithm proposed in this study. Thick line on the maps delineates the 50% isopleth of multi-year ice

5 fraction.

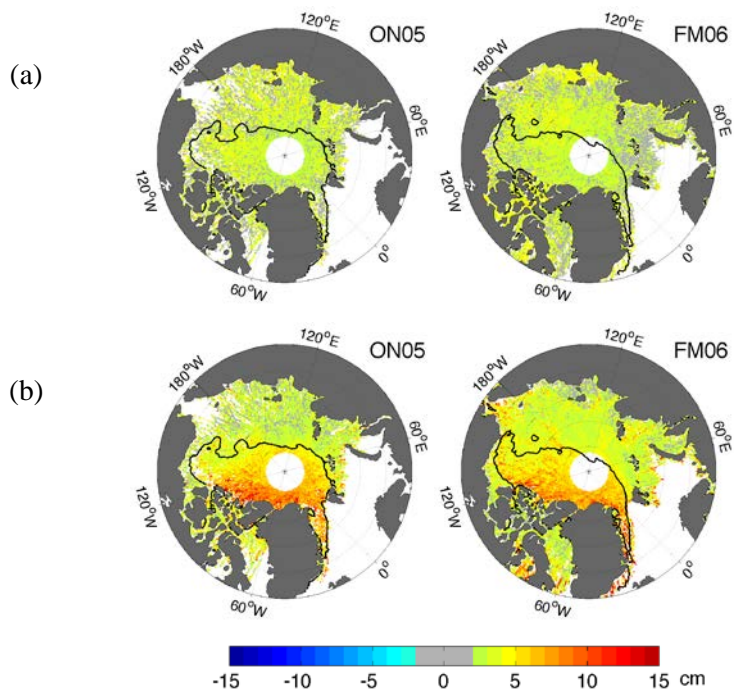
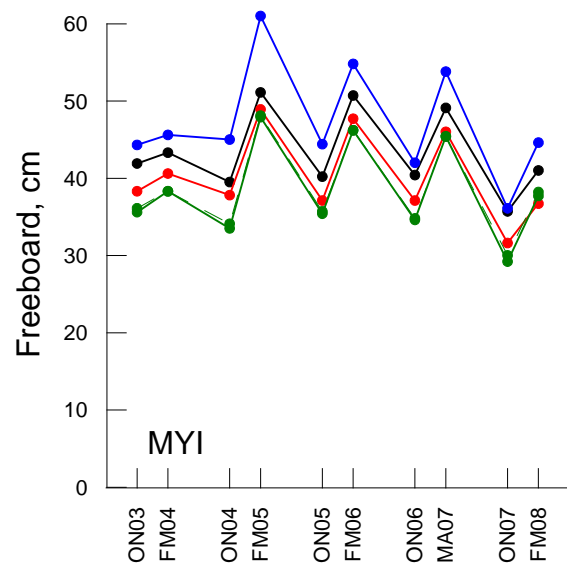
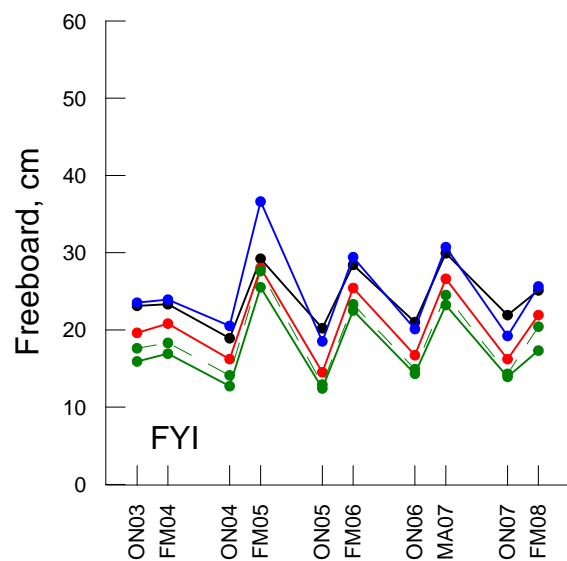
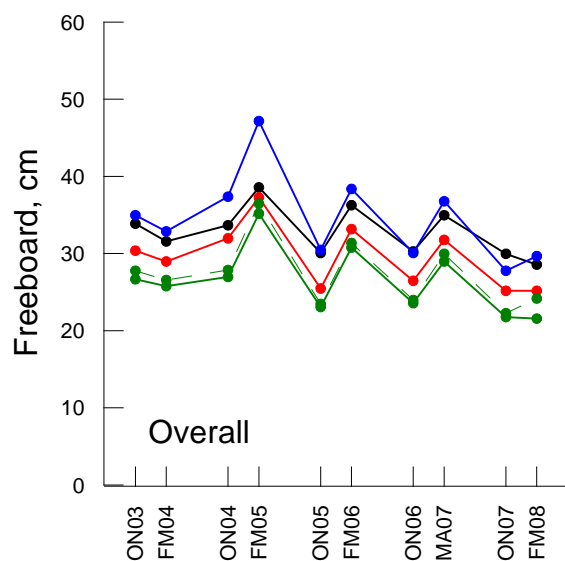


Figure 98. Adjustments of freeboard retrieved using the TP method accounting for (a) snow depth on top of new ice in leads and (b) leads width with respect to ICESat footprint area for the periods ON05 and FM06. These adjustments are estimated following the methods described in Kwok and Cunningham (2008) and Kwok et al. (2009).



- GSFC
- Lowest level elevation, 25 km
- Tiepoints
- - Tiepoints modified
- Tiepoints modified, adjustments applied

Figure 109. Freeboard time series calculated over the whole arctic basin, FYI and MYI areas as provided in the GSFC product (black) and retrieved in this study by the LLE method when using shorter along-track averaging scales (red), the TP method (green), the TP method, which includes the improvements in the freeboard retrieval algorithm proposed in this study without (green dashed) or with (blue) adjustments for snow depth and leads width applied.