Assessing Scale Dependence on Local Sea Level Retrievals from Laser Altimetry Data over Sea Ice

Liuxi Tian, Hongjie Xie *, Stephen F. Ackley and Alberto M. Mestas-Nuñez

NASA MIRO Center for Advanced Measurements in Extreme Environments and Department of
Geological Sciences, University of Texas at San Antonio, San Antonio, TX 78249, USA;
liuxi.tian@my.utsa.edu (L.T.); stephen.ackley@utsa.edu (S.F.A.); alberto.mestas@utsa.edu (A.M.M.-N.)
* Correspondence: Hongjie.Xie@utsa.edu; Tel.: +1-210-4585445

check for
updates

Abstract: The measurement of sea ice elevation above sea level or the "freeboard" depends upon an accurate retrieval of the local sea level. The local sea level has been previously retrieved from altimetry data alone by the lowest elevation method, where the percentage of the lowest elevations over a particular segment length scale was used. Here, we provide an evaluation of the scale dependence on these local sea level retrievals using data from NASA Operation IceBridge (OIB) which took place in the Ross Sea in 2013. This is a unique dataset of laser altimeter measurements over five tracks from the Airborne Topographic Mapper (ATM), with coincidently high-spatial resolution images from the Digital Mapping System (DMS), that allows for an independent sea level validation. The local sea level is first calculated by using the mean elevation of ATM L1B data over leads identified by using the corresponding DMS imagery. The resulting local sea level reference is then used as ground truth to validate the local sea levels retrieved from ATM L2 by using nine different percentages of the lowest elevation (0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4%) at seven different segment length scales (1, 5, 10, 15, 20, 25, and 50 km) for each of the five ATM tracks. The closeness to the 1:1 line, R^2, and root mean square error (RMSE) is used to quantify the accuracy of the retrievals. It is found that all linear least square fits are statistically significant ($p < 0.05$) using an F test at every scale for all tested data. In general, the sea level retrievals are farther away from the 1:1 line when the segment length scale increases from 1 or 5 to 50 km. We find that the retrieval accuracy is affected more by the segment length scale than the percentage scale. Based on our results, most retrievals underestimate the local sea level; the longer the segment length (from 1 to 50 km) used, especially at small percentage scales, the larger the error tends to be. The best local sea level based on a higher R^2 and smaller RMSE for all the tracks combined is retrieved by using 0.1–2% of the lowest elevations at the 1–5 km segment lengths.

Keywords: lowest elevation method; segment length; Operation IceBridge; Ross Sea; Antarctica

1. Introduction

Sea ice is an important climate factor and indicator [1,2]. Compared to sea ice area and extent which have been measured by passive microwave remote sensing for more than 40 years [3–5], the ice thickness measurement technique from satellites is still under development and faces many challenges. A key area of interest for scientists is the use of radar and laser altimetry data to extract ice or total freeboard, which is then used to derive sea ice thickness [6–9]. Total freeboard is the height of the snow depth and ice surface above the local sea level (open water or very thin sea ice, <5 to 10 cm). While radar signals may be reflected from somewhere within the snow layer rather than the snow–ice interface [10], laser signals usually return from the air–snow interface, so the total freeboard can be

derived more accurately from laser altimetry [11]. To derive the freeboard, an accurate local sea level reference is needed.

To determine the local sea level, the precise identification of leads (open water, thin ice) is crucial. As the elevation of open water or thin sea ice is lower than that of the nearby sea ice/snow-covered surface, a certain percentage of the lowest elevation (footprints) along a range of a flight track is averaged to get an estimate of the local sea level [12,13], the so-called lowest elevation method. ICESat depends on this method to derive the freeboard. When using the lowest elevation method, there are two key factors affecting the derived local sea level: the length of the segment range and the percentage of the lowest elevations that are chosen. Zwally et al. (2008) [12] assumed that a local sea surface could be represented by the lowest 2% elevation of ICESat laser shots, within a 50-km segment length. Price et al. [14] assumed that the sea surface may be represented by the mean of the lowest 5% of elevation measurements along each track in their small-scale study. Since freeboard estimates from altimetry data such as ICESat depend on the lowest elevation method to retrieve the local sea level, the purpose of this study is to assess the effects of different segment length scales and percentage elevation scales on local sea level retrieval.

To assess the effects of different segment lengths or percentage elevations on local sea level retrieval is difficult due to the lack of sea level tie point validations. The NASA Operation IceBridge (OIB) program, however, provides a unique dataset with the acquisition of laser altimeter measurements and coincident high-spatial resolution optical imagery [15], to allow for a sea level tie point validation [16]. In this study, we assess the effect of different segment length and percentage elevation scales on the lowest elevation method by using the OIB dataset. We take the Ross Sea OIB campaign (November 2013) as our study location.

2. Data and Methods

2.1. Data

We use data collected as part of the OIB mission during the 2013 field campaigns over the Ross Sea, including laser altimeter (Airborne Topographic Mapper—ATM) and high-resolution optical imagery (Digital Mapping System—DMS) (Figure 1). ATM data are surface elevation data acquired by an airborne light detection and ranging (LIDAR) system using a 532 nm wavelength laser beam. The ATM L1B data, at a typical altitude of 500 m above ground level, a laser pulse rate of 5 kHz, and a scan width of 22.5 degrees off-nadir, have nominal spatial resolution of 1 m [17]. The ATM L2 data are the resampled data at the distance interval averaging 0.5 s (approximately 60 m) worth of data along the flight track and a fixed 80 m across-track nadir platelet as well as three or five additional platelets that together span the entire swath of the ATM scan [18]. The DMS data are high-resolution natural color and panchromatic images with resolution ranging from 0.015 to 2.5 m depending on flight altitude (0.1 m at altitude of 457 m). The OIB data coordinates are referenced to the WGS84 ellipsoid [19]. As the sea surface height (SSH) is influenced by the geoid, tidal forces, atmospheric pressure, and ocean dynamic topography, all the data are corrected by subtracting the DTU15 mean sea surface (MSS) [20]. The DTU15 MSS includes geoid undulation and ocean mean dynamic topography, the largest variations of SSH. We obtain sea surface height anomalies (SSHAs) by using SSHs (i.e., the ATM elevation) minus the DTU15 MSS. These SSHAs are used to calculate the local sea level in this study.
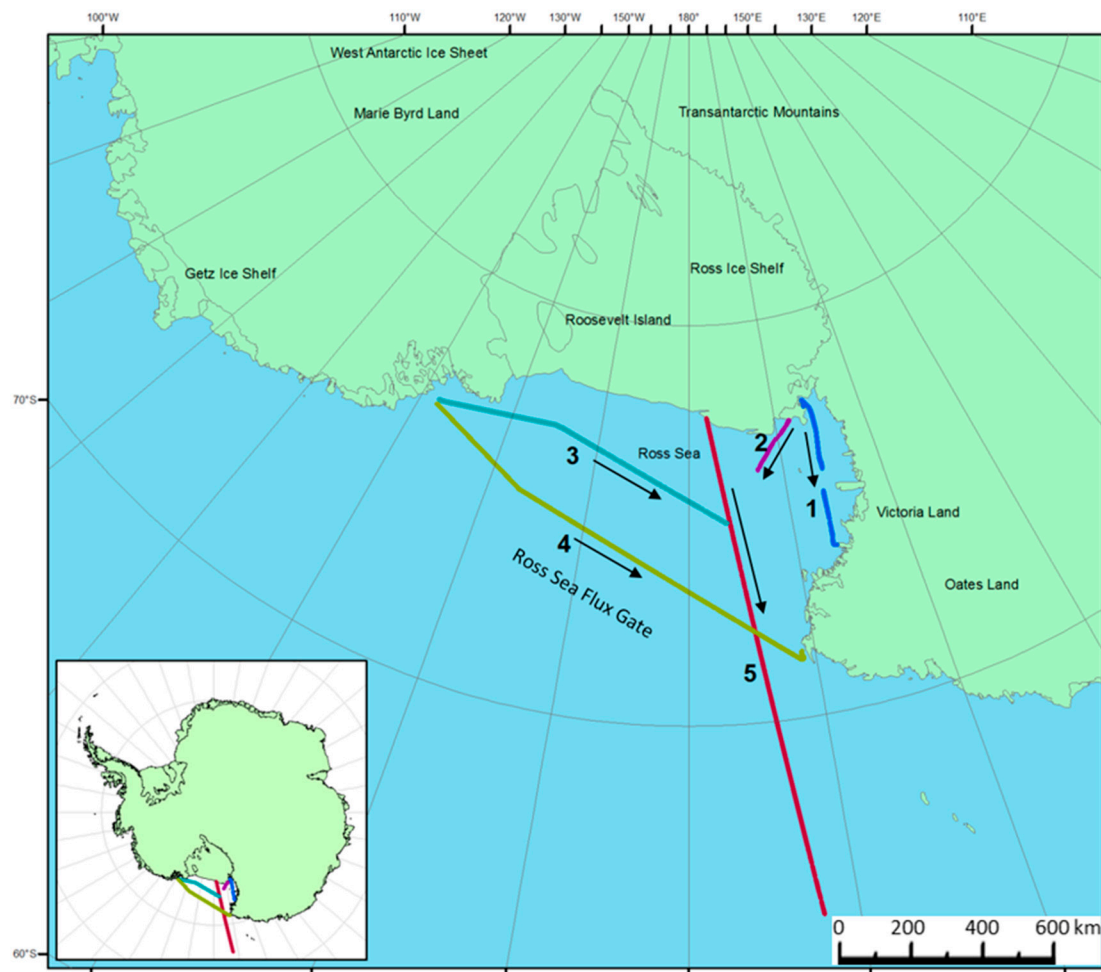
Figure 1. IceBridge flight tracks in the Ross Sea, Antarctica, on November 20 and 27, 2013. The arrows show the direction of each track that is used for data analysis and discussion purposes (adapted from [21]).

2.2. Local Sea Level Retrieved from ATM with Leads Identified by Using DMS

DMS images are used to detect the leads, which are then used to retrieve local sea levels from the ATM elevations over these leads. We assume that these local sea levels are ground truth as the DMS (optical images) could exactly provide the lead locations. As leads are optically dark features with low reflected pixel intensities (brightness values), while snow/sea ice come with high reflected intensities, in the high-resolution natural color DMS images, the distribution modes of the brightness values provide a practical way to separate leads from sea ice [22]. These retrieved local sea levels from ATM L1B are then treated as ground truth and used to validate the local sea level retrievals by using the lowest elevation method from ATM L2 data alone.

2.3. Local Sea Level Retrievals from the Lowest Elevation Method

The lowest elevation method is based on assuming that a certain percent of the lowest SSHAs along a fixed segment of a flight track can define the local sea level. In this study, we use different percentage and segment length scales to retrieve the local sea level from ATM L2 and compare them with the "ground truth" local sea level when DMS images are involved, as explained in Section 2.2. We use a total of 63 combinations coming from 7 different segment lengths (1, 5, 10, 15, 20, 25, and 50 km) and 9 different percentage elevation values (0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4%) for each track. We also resample the ATM L2 data with a 170 m gap, simulating the spacing of ICESat data, and repeat the same processes to obtain the local sea levels.

To summarize the comparisons between the local sea level from the lowest elevation method and the ground truth for the different scale combinations, linear least square fits to each scatter plot comparison are used and the significance of these fits is verified using F statistics. To assess the strength of these comparisons, the coefficients of determination (R^2) (Equation (1)) and the root mean square error (RMSE) (Equation (2)) are used. The closer the fitting line is to 1:1, the closer the R^2 is to 1, and the smaller the RMSE is to 0 m define the best estimation.

$$R^2 = \left(\frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{\left[n \sum x^2 - (\sum x)^2\right]\left[n \sum y^2 - (\sum y)^2\right]}} \right)^2 \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - x_i)^2}{n}} \tag{2}$$

where x represents the ground truth value, y represents the estimated value from the lowest elevation method, i represents one tie point, and n stands for the number of tie points.

3. Results

3.1. Local Sea Level from ATM L1B and DMS Images (Ground Truth)

The local sea levels retrieved from ATM L1B and DMS images for the five tracks are shown in Figure 2 and are treated as the ground truth of this study to validate the local sea levels retrieved from the lowest elevation method. Each dot represents the mean elevation of ATM L1B over leads identified by using the corresponding DMS imagery. The standard deviation quantifies how close the elevations of ATM L1B values over leads of one DMS image are to the mean elevation value. Track 3 shows the flattest local sea levels among the five tracks, also with small standard deviations (red). The local sea levels vary the most along track 4. Track 5 has the highest local sea level with the largest standard deviation, especially over the inner part of the fluxgate (0 to 400km) (Figure 1). The local sea level variations are affected by many factors including winds, currents, variations in gravity, and temperature.
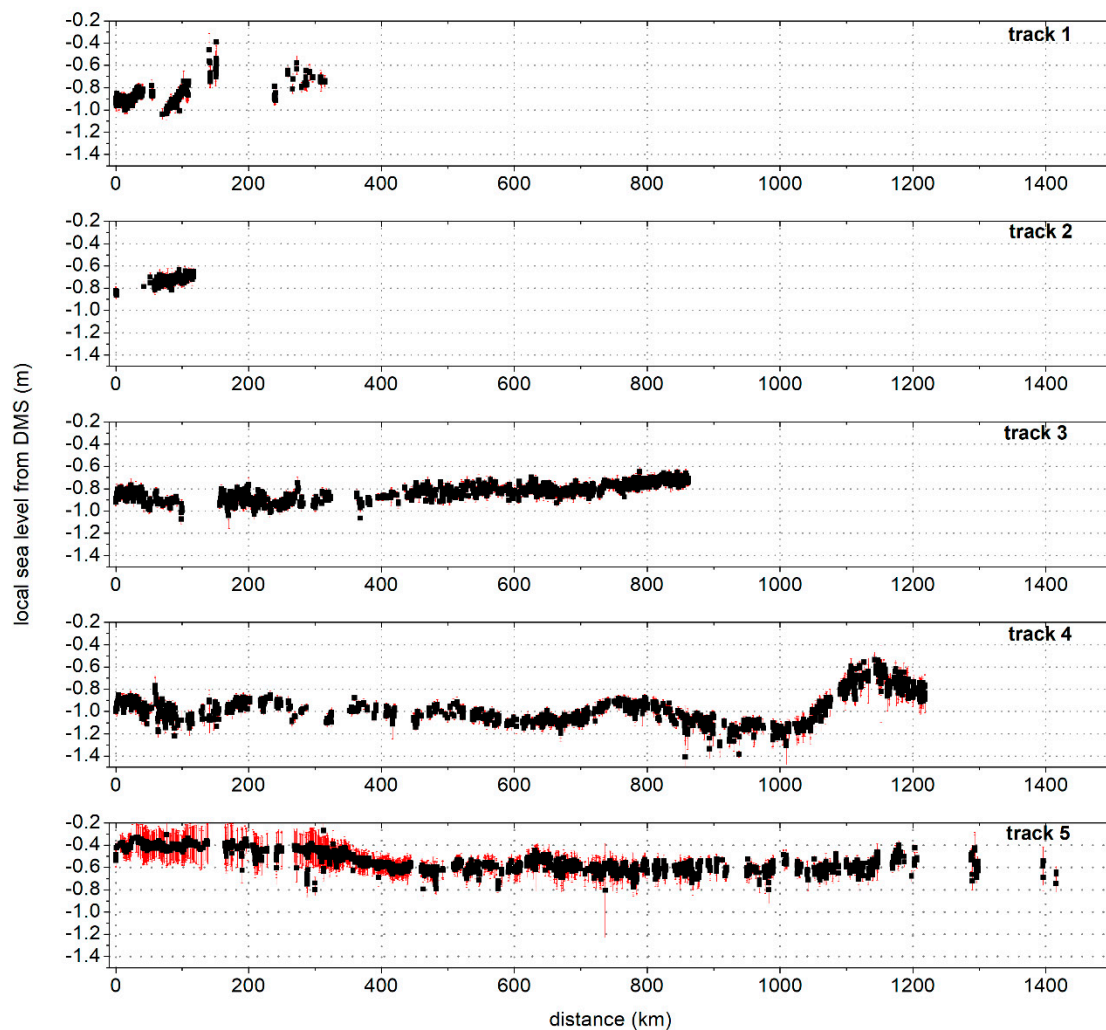
Figure 2. Local sea level with ±1 standard deviation (red) for each track by using the Digital Mapping System's (DMS) lead detection. The direction in which the along-track distance is measured (start and end) is indicated with the arrows in Figure 1 (adapted from [21]).

3.2. Linear Least Square Fits at Different Scale Combinations

Figures 3–7 show linear fits with the 63 scale combinations for each track. There are a total of nine panels (percentage scale) for each track, and each panel has seven fitting lines (segment length scales). All linear least square fits with each segment length scale and percentage scale for all five tracks are statistically significant with greater than 95% confidence using an F test. For tracks 1 and 2 (Figures 3 and 4), it is found that most lines appear to lie below the 1:1 line except the three 1 km lines above the 1:1 line, where the percentage scale is 3%, 3.5 %, or 4%, of track 1 and some lower parts of the 1 km lines of track 2. This indicates local sea levels from the lowest elevation method are almost all underestimated except some 1 km conditions that are closer or slightly overestimated for tracks 1 and 2. This suggests for the total freeboard found that by subtracting its elevation from the local sea level, elevation will be overestimated, except at the 1 km segment length scale. It is also found that for each panel, the local sea level retrievals are farther away from the 1:1 line when the segment length scale increases from 1 to 50 km; however, when the percentage scale is large (3%, 3.5 %, or 4%), the fitting lines at the 1 km segment length are farther away from the 1:1 line than the 5 km segment length of track 1.

For track 3 (Figure 5), all lines are found to lie below the 1:1 line except some lower local sea level parts. This means that local sea levels from the lowest elevation method are mostly underestimated for

track 3. It also indicates that for each panel, the sea level estimations are farther away from the 1:1 line when the segment length scale increases from 1 to 50 km.

For track 4 (Figure 6), we see all lines appear to lie below the 1:1 line except the 1 km lines that are very slightly above the 1:1 line. This indicates local sea levels from the lowest elevation method are all underestimated except for the 1 km conditions. Similar to the other tracks, the sea level retrievals are farther away from the 1:1 line as the segment length scale increases from 5 to 50 km.

For track 5 (Figure 7), it shows all lines lie partly above the 1:1 line (lower local sea level section) and partly below the 1:1 line (higher local sea level section). This indicates that local sea levels retrieved from the lowest elevation method for track 5 are very complex and the values are overestimated where there are lower local sea levels (outer part of fluxgate >400 km) and underestimated where there are higher local sea levels (inner part of the fluxgate < 400 km).

In general, the sea level retrievals are farther away from the 1:1 line when the segment length scale increases from 1 or 5 to 50 km. The nine panels of each track are very similar to each other. This means that the percentage scale does not make a difference to the results for all tracks. For tracks 3, 4, and 5, all lines for each panel of these three tracks are much closer to each other than tracks 1 and 2, indicating less impact by the segment length scale for tracks 3, 4, and 5 than for tracks 1 and 2. Track 5 shows the most overestimated local sea levels, especially further northwards (>400 km).
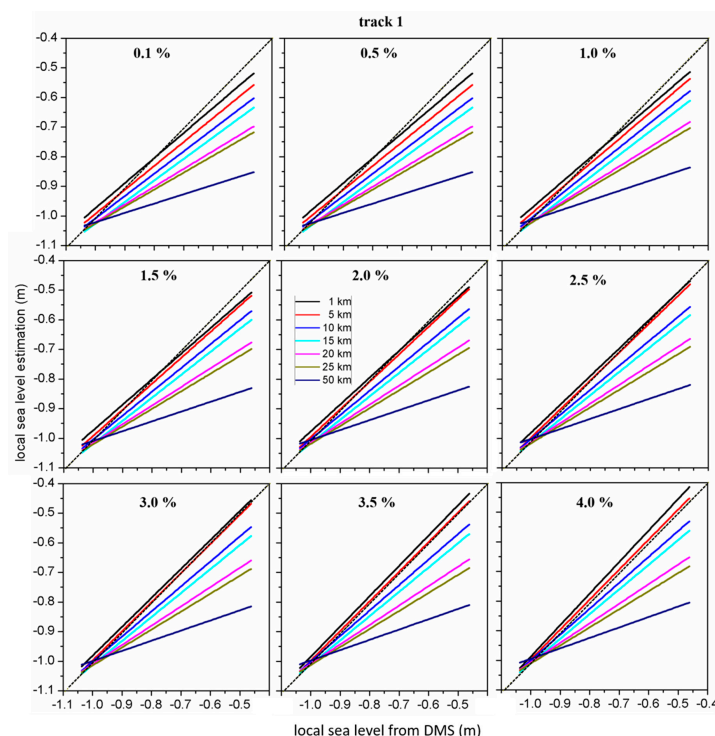

Figure 3. Linear fit lines of track 1 between retrievals from the lowest elevation method and those from the DMS method (ground truth) at 7 segment length scales: 1, 5, 10, 15, 20, 25, and 50 km, and 9 percentage scales: 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4%. The dashed line is the 1:1 line. All linear least square fits are statistically significant with greater than 95% confidence using an F test.
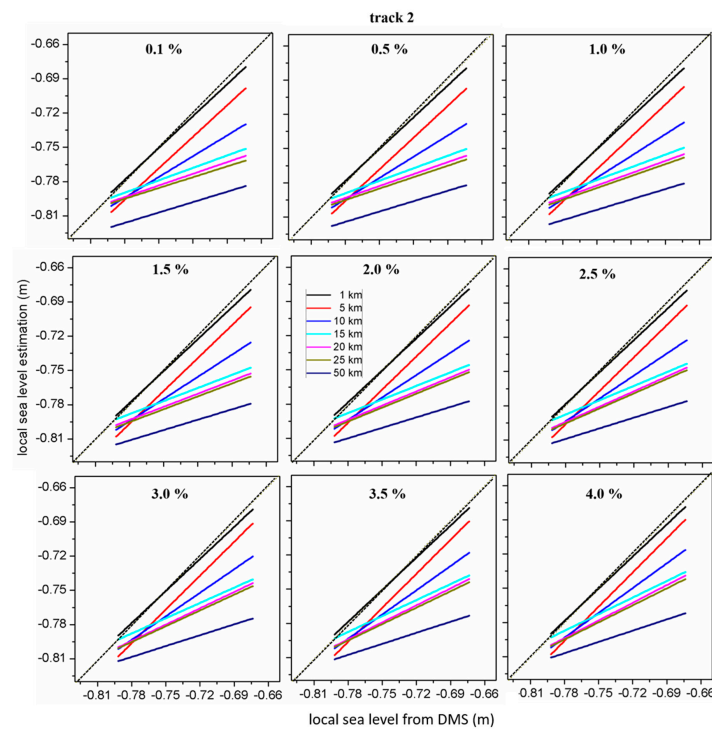
Figure 4. Linear fit lines of track 2 between retrievals from the lowest elevation method and those from the DMS method (ground truth) at 7 segment length scales: 1, 5, 10, 15, 20, 25, and 50 km, and 9 percentage scales: 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4%. The dashed line is the 1:1 line. All linear least square fits are statistically significant with greater than 95% confidence using an F test.
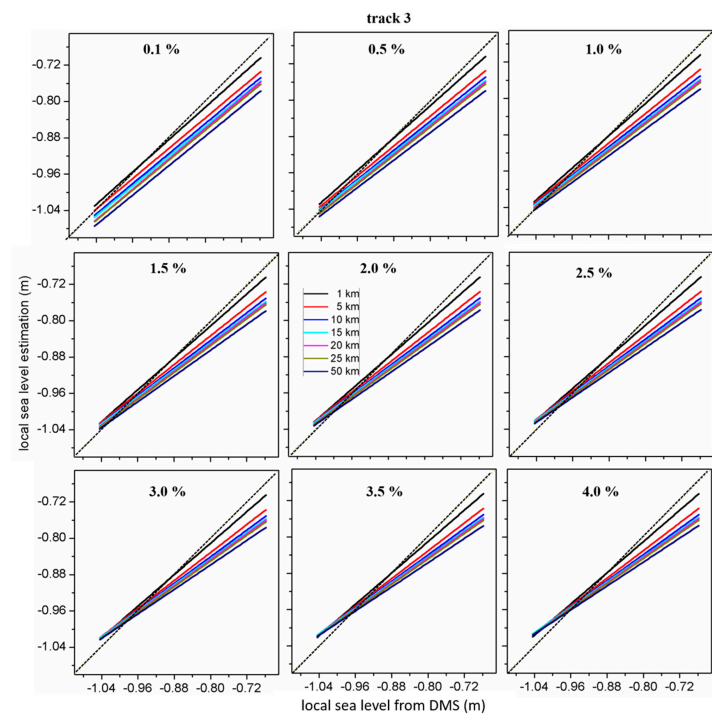

Figure 5. Linear fit lines of track 3 between retrievals from the lowest elevation method and those from the DMS method (ground truth) at 7 segment length scales: 1, 5, 10, 15, 20, 25, and 50 km, and 9 percentage scales: 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4%. The dashed line is the 1:1 line. All linear least square fits are statistically significant with greater than 95% confidence using an F test.
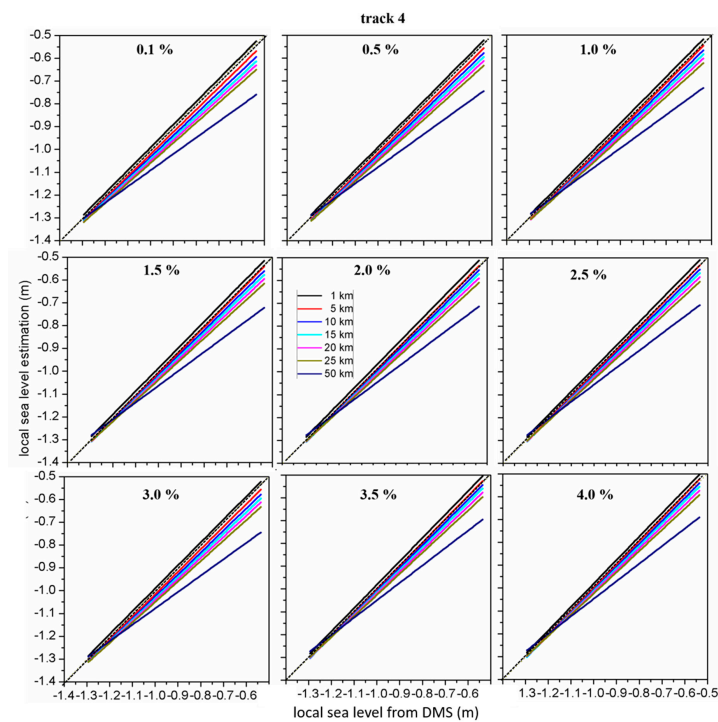
Figure 6. Linear fit lines of track 4 between retrievals from the lowest elevation method and those from the DMS method (ground truth) at 7 segment length scales: 1, 5, 10, 15, 20, 25, and 50 km, and 9 percentage scales: 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4%. The dashed line is the 1:1 line. All linear least square fits are statistically significant with greater than 95% confidence using an F test.

Figure 7. Linear fit lines of track 5 between retrievals from the lowest elevation method and those from the DMS method (ground truth) at 7 segment length scales: 1, 5, 10, 15, 20, 25, and 50 km, and 9 percentage scales: 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4%. The dashed line is the 1:1 line. All linear least square fits are statistically significant with greater than 95% confidence using an F test.
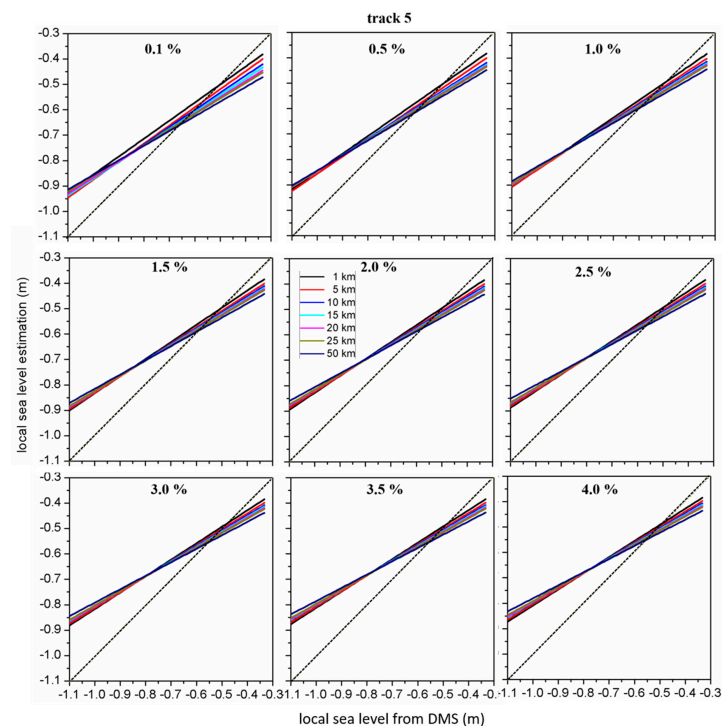
3.3. Quantifying the Accuracy of Local Sea Level Retrievals

Figure 8 shows the R^2 and root mean square error (RMSE) for every track. For track 1, the R^2 values range from 0.28 to 0.82, and RMSE values range from 0.04 to 0.15 m. As shown, the R^2s are the largest at 1, 5, and 10 km, followed by 15, 20, 25, and 50 km, when the percentage scales are from 0.1% to 1.5%. The RMSEs are the smallest at 1 km, followed by 5, 10, 15, 20, 25, and 50 km, when the percentage scales are from 0.1% to 2.0%. When the percentage scales are from 2.5% to 4.0%, the 10 km fitting line has the largest R^2 and lowest RMSE. Overall, combining R^2 and RMSE results, the 1 km segment length with the percentage scales from 0.1% to 1.5 % gives the best estimations.
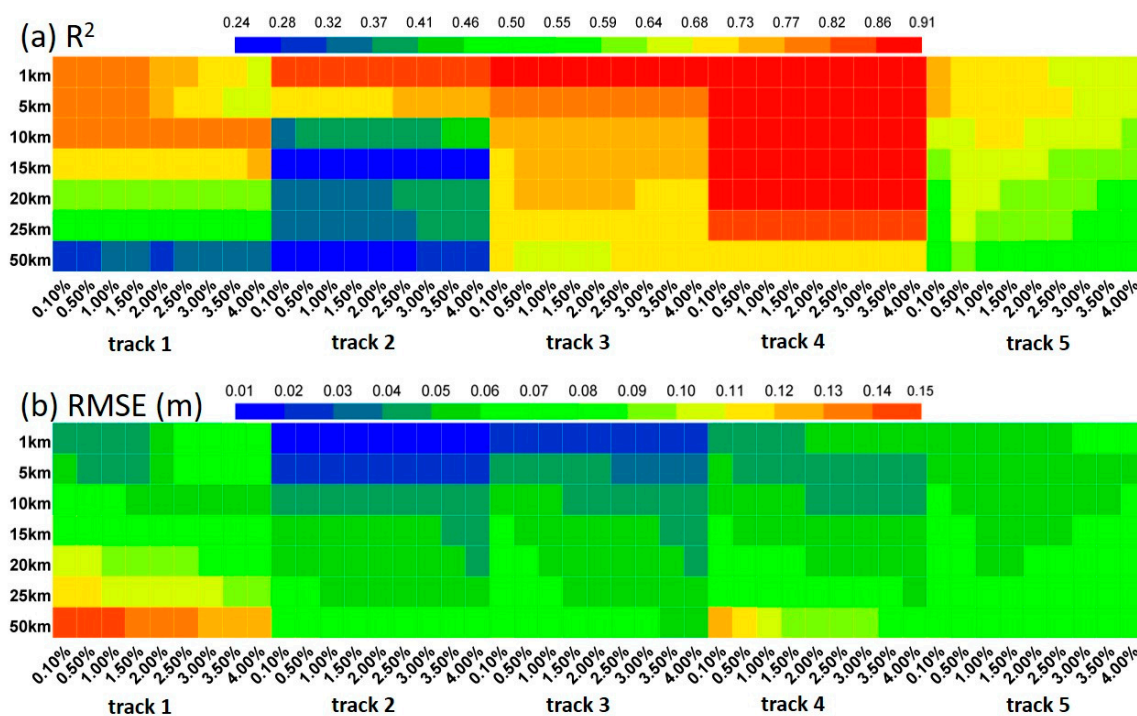

Figure 8. Coefficient of determination (R^2) (**a**) and root mean square error (RMSE (m)) (**b**) of all tracks between retrievals from the lowest elevation method and retrievals from the DMS method (ground truth) at 7 segment length scales: 1, 5, 10, 15, 20, 25, and 50 km, for 9 percentage scales: 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4% (the x-axis).

For track 2, the R^2 values range from 0.24 to 0.86, and RMSE values range from 0.01 to 0.10 m. As shown, the R^2s are the largest at 1 km, followed by 5 and 10 km. The 15 km fitting line has the smallest R^2. The RMSEs are the smallest at 1 km, followed by 5, 10, 15, 20, 25, and 50 km. Overall, combining R^2 and RMSE results, the 1 km segment length fitting lines with any percentage scale give the best retrievals.

For track 3, the R^2 values range from 0.64 to 0.91, and RMSE values range from 0.02 to 0.10 m. As shown, the R^2s are the largest for 1 km segment lengths, followed by 5, 10, 15, 20, 25, and 50 km. The RMSEs are the smallest for 1 km segment lengths, followed by 5, 10, 15, 20, and 25 km. Overall, the 1 km segment length fitting lines with any percentage scale give the best retrievals.

For track 4, the R^2 values range from 0.68 to 0.91, and RMSE values range from 0.04 to 0.13 m. As shown, the R^2s are the largest at 1, 5, 10, 15, and 20 km segment lengths, followed by 25 and 50 km. The RMSEs are the smallest for 1 km segment lengths, followed by 5, 10, 15, 20, 25, and 50 km, when the percentage scales are from 0.1% to 1.5%. Overall, the 1 km segment length fitting lines with the percentage scales from 0.1% to 1.5 % give the best estimations.

For track 5, the R^2 values range from 0.46 to 0.77, and RMSE values range from 0.05 to 0.10 m. As shown, the R^2s are the largest for 1 and 5 km segment lengths, followed by 10, 15, 20, 25, and 50 km,

when the percentage scales are from 0.1% to 2.0%. The RMSEs are the smallest for 1 and 5 km segment lengths, followed by 10, 15, 20, 25, and 50 km, when the percentage scales are from 0.1% to 2.0%. Overall, 1 and 5 km segment length fitting lines with the percentage scale of 0.1% give the best retrievals.

Since the track length and thus the amount of data vary greatly from track to track, and to generate more robust statistics, we computed the R^2 and RMSE for all tracks combined and show these results in Figure 9. The overall R^2 is greater than 0.9 for all segment lengths and percentage scale combinations with maximum values (>0.95) for the 1–5 km track segments and percentage scales between 0.1% and 2%. Based on the RMSE, the best agreement (RMSE < 0.04) is also found for 1–5 km track segments for percentage scales between 0.1% and 2%. To estimate the impact of using coarser scales to estimate the local sea level, we used all the tracks combined at 0.1% and 1 km resolutions and computed the bias (mean difference), which resulted in 0.004 m. The same bias calculation using the coarser 2% and 50 km scales as in [21] gives −0.045 m, indicating a total local sea level bias change of about −0.05 m or an underestimation of the local sea level by the coarser scales relative to the finer ones. Thus, the freeboard estimates in [21] may have been overestimated by about 0.05 m. In summary, based on the overall comparison, 1–5 km track segments have the best retrievals for 0.1–2 percentage scales based on both R^2 and RMSE criteria.
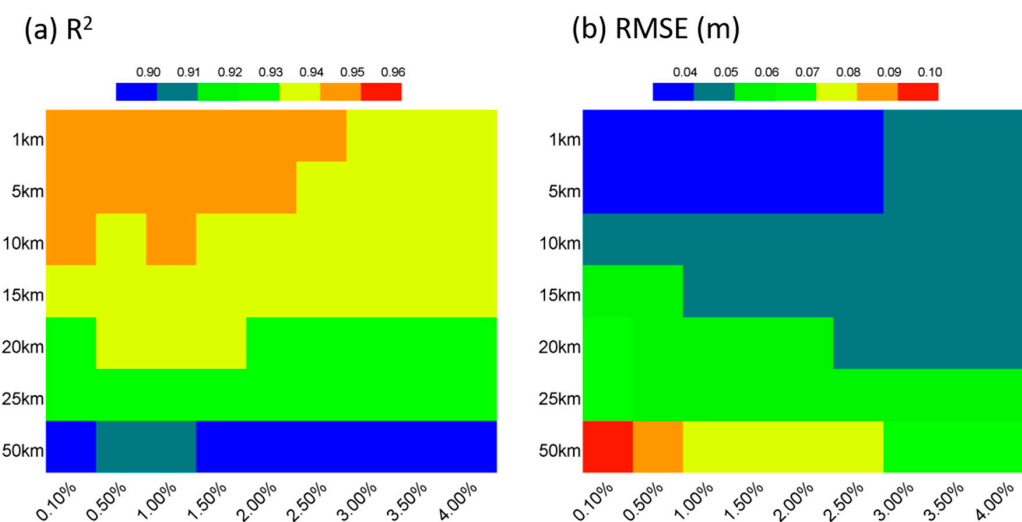

Figure 9. The R^2 (**a**) and RMSE (m) (**b**) for all the tracks combined between retrievals from the lowest elevation method and retrievals from the DMS method (ground truth) at 7 segment length scales: 1, 5, 10, 15, 20, 25, and 50 km, for 9 percentage scales: 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4% (the x-axis).

4. Discussion

In a previous study, Wang et al. (2013) [16] evaluated local sea level estimation by using the lowest 0.1%, 0.2%, 0.5%, 1%, and 2% of all ATM L1B elevations within each 1 km range. They based their study on four ~30 km sections of the Oct. 21, 2009 flight line over the Bellingshausen–Amundsen Seas, where one DMS image per 1 km length was used to manually select open leads and/or thin ice, with the corresponding ATM elevations retrieved as ground truth. They found that the lowest 0.2% of ATM elevations matched well with the ground truth. However, they did not apply the segment length scale [16]. In this study, we found that the percentage scale is not a determining factor as compared to the segment length scale. Considering that the flight lines used in the Wang et al. (2013) [16] study are much shorter than in this study, this study provides a more complete picture by evaluating the influences of both the percentage scale and segment length scale.

Some studies retrieved the local sea level from laser altimetry by assuming the sea surface to be represented by the mean of the lowest empirically determined percentage of elevation measurements

along an empirically determined segment length track, without any or sufficient ground truth data to access and evaluate the fitness of such retrievals [6,12,14,23]. In this study, we take advantage of the optical images of the DMS along with altimetry ATM data to investigate the accuracy of the retrievals that result from using different scale combinations. We find that the retrieval accuracy is affected more by the segment length scale than the percentage scale. Based on our results, most retrievals underestimate the local sea level; the longer the segment length from 1 to 50 km used, especially when the percentage scale is small, the larger the error tends to be. For track 5, the retrievals show overestimates for locations where the local sea level is lower (outer part of fluxgate (>400 km)) and underestimates where the local sea level is higher (inner part of the fluxgate (<400 km)). There are potentially two reasons for this phenomenon. One is because the sea levels have considerable variations over longer-distance scales (Figure 2). A longer segment length would have larger variations of the local sea level; therefore, longer segment lengths will potentially result in higher errors. The local sea level variations are affected by many factors including winds, currents, variations in gravity, and temperature. Since these processes are length scale-dependent, longer segments would include more uncertainties related to these processes. The second reason is because if the opening water/leads are not wide enough, the higher percentages may mix nearby ice elevation with leads. This could happen at all length scales if the ice is compact or the ice concentration is 90% and above. However, short segment lengths with a lower percentage may not collect enough open water/leads data. For track 5, as new ice produced in coastal polynyas is transported northward by katabatic winds off the ice shelf, ice is overall thicker and tends to compact beyond the fluxgate (Figure 1). Longer segment lengths could mix ice elevations with open water/thin ice elevations and therefore overestimate the local sea levels. So, the overestimation of track 5 could be attributed to mixed ice elevations and open water/thin ice elevation where the ice is more compact due to the ice being pushed away from the inner part of the continental shelf, caused by katabatic winds off the Ross Ice Shelf.

Kern and Spreen (2015) [23] used the ICESat data to calculate the local sea level. They used the 2% lowest elevation at a 50 km segment length. ICESat has a 70-m footprint size and a 172 m gap sampling. The ATM L2 data (footprint of 60 x 80 m) therefore have a similar footprint to ICESat. To explore the effect of the 172 m gap between footprints, we resampled the ATM L2 data with a 170 m gap and took the same tracks of ICESat (using the data of 2005 as an example) and compared the count number (Table 1). The result indicates that the 170 m gap of ATM L2 has, understandably, less counts than the original data, but similar to the ICESat 2005 counts. Figure 10 computes the simulated gap data, and comparing with the full dataset used in Figure 8, the results indicate, in general, the trend along the percentage scale in Figure 10 becoming flatter than the data without the gap in Figure 8. This means that the gap causes the percentage scale to have less effect. We compare the indices R^2 and RMSE (m) for each track with a 1 km segment length by a 0.1% scale between OIB with and without a 170 m gap. As shown in Table 2, it is indicated that fits with a 170 m gap are better for tracks 1 and 2, while these are worse for tracks 3, 4, and 5 (Table 2). As high-resolution optical imagery is not available simultaneously for altimetry data like ICESat, our study provides useful information for the optimal selection of segment length and percentage scales, for applying the lowest elevation method to derive the local sea level.

Table 1. Comparison of count number among Operation IceBridge (OIB) with and without a 170 m gap and ICESat data (taking 2005 as an example).

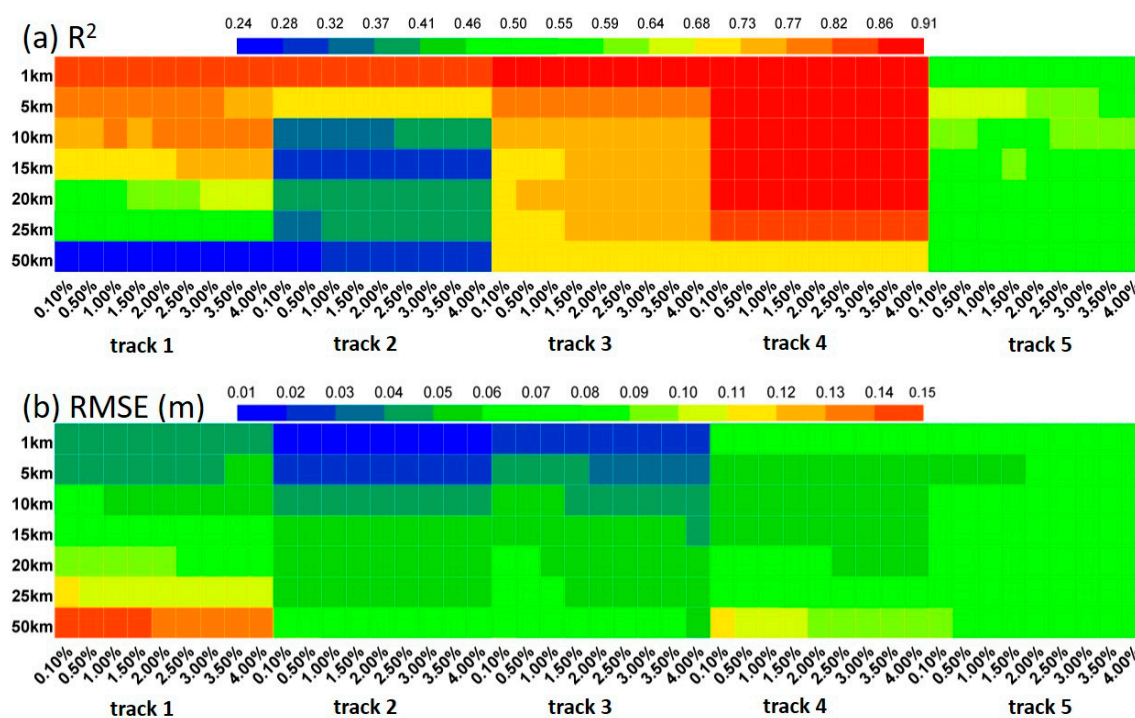| Tracks	Count		
	OIB	**OIB with 170 m gap**	**ICESat (2005)**
track 1	22,236	1458	2458
track 2	3077	312	1213
track 3	51,364	3338	10,329
track 4	112,108	6313	11,548
track 5	125,531	7319	12,547

Figure 10. R^2 (**a**) and RMSE (m) (**b**) for all tracks of ATM L2 170 m gap data between local sea level retrievals from the lowest elevation method and retrievals from the DMS method (ground truth) at 7 segment length scales: 1, 5, 10, 15, 20, 25, and 50 km, for 9 percentage scales: 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4%.

Table 2. Comparison of the indices R^2 and RMSE(m) for each track with a 1 km segment length and 0.1% scale between OIB with and without a 170 m gap.

Tracks	OIB		OIB with 170 m gap	
	R^2	RMSE(m)	R^2	RMSE(m)
track 1	0.8066	0.0457	0.8207	0.0410
track 2	0.8204	0.0140	0.8332	0.0136
track 3	0.9058	0.0231	0.8876	0.0237
track 4	0.8956	0.0471	0.8632	0.0641
track 5	0.7295	0.0517	0.5158	0.0753

Since the reflectivity of laser shots on thick sea ice is much higher than on leads, this is another way to classify the leads. The mean reflectivity values of leads for all 2013 IceBridge tracks in the Ross Sea are less than 0.25, mostly around or less than 0.15, with a standard deviation of 0.05 [21]. Kwok and others (2012) [24] estimated the local sea level (tie points) along 250 m segments for Arctic sea ice conditions with a reflectivity value (R) of 0.25 as the threshold for L1B data to extract lead shots. In this study, we did not consider the effect of reflectivity and only took the Ross Sea data as an example to assess the estimation, so more research about the influence of the reflectivity threshold is needed.

The underestimated local sea levels will result in overestimating the total freeboard and the overestimated local sea levels will result in underestimating the total freeboard. Our results provide the reference to derive the local sea level using the lowest elevation method especially for those altimetry data such as ICESat without simultaneous high-resolution imagery data. As for ICESat 2, it actually uses a totally different method to retrieve the local sea level, due to its small footprint size (17m) and 70 cm interval, compared to 70 m and the 170 m interval for ICESat [25]

5. Conclusions

The accurate retrieval of the total freeboard from laser altimetry over sea ice is based on the accurate retrieval of the local sea level which depends on the accurate identification of open water and thin ice. For some altimetry data such as ICESat, to obtain an appropriate segment length or percentage of the lowest elevations is difficult because of the lack of sea level tie point validation. It is feasible to access the scales of segment length or lowest elevation percentage by using OIB data because of the corresponding DMS and ATM data. The local sea level is first calculated by using the mean elevation of ATM L1B over leads identified by using the corresponding DMS imagery. This sea level reference is then used as ground truth to validate sea level retrievals from ATM L2 at seven different segment length scales (1, 5, 10, 15, 20, 25, and 50 km) and nine percentage scales (0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, and 4%). The R^2 and root mean square error (RMSE) are used to quantify the accuracy of the retrievals.

It is found that all linear least square fits are statistically significant with greater than 95% confidence using an F test at every scale for all tracks ($p < 0.05$). For the five tracks individually at the 63 possible scale combinations, R^2s range from 0.24 to 0.91, and RMSEs range from 0.01 to 0.15 m. It is found segment length scales are more influential than percentage elevation scales on the retrievals. The sea level retrievals are generally farther away from the 1:1 line when the segment length scale increases from 1 or 5 to 50 km. Most retrievals are underestimated, which will result in overestimating the total freeboard. However, most local sea level retrievals for track 5 at the outer part (>400 km) of the fluxgate are overestimated, which will result in underestimating the total freeboard. This local sea level overestimation can be attributed to mixed ice elevations and open water/thin ice elevation where the ice is more compact due to the ice being pushed away from the inner part of the continental shelf, caused by katabatic winds off the Ross Ice Shelf. For tracks 3, 4, and 5, all lines for each panel of these three tracks are much closer to each other than tracks 1 and 2, indicating less impact by the segment length scale for tracks 3, 4, and 5 than for tracks 1 and 2. The 1 km segment length by the 0.1% scale has the best retrievals overall for all tracks. At these scales, R^2s are within the range from 0.73 to 0.91, and RMSEs are within the range from 0.01 to 0.05 m.

When ATM L2 is resampled to the 170 m gap simulating ICESat data, the results show a similar pattern and trend among the segment length scales, but the trend among the percentage scales becomes flatter than the data without the gap. This means that the gap results in even less impact from the percentage scale. Using the 1 km segment length and 0.1% scale between OIB with and without a 170 m gap indicates that fits with a 170 m gap are better in tracks 1 and 2 and worse in tracks 3, 4, and 5 than using the data without a gap.

Combining all five tracks to estimate overall statistics for the 63 possible combinations of percentage and track length scales, we find that R^2 and RMSE vary between 0.9 and 0.96 and 0.04 and 0.1, respectively. Both heat maps give favorable values with $R^2 > 0.94$ and RMSE < 0.04 m for percentage scales in the 0.1–2% range and track segments in the 1–5 km range. Using the 0.1% and 1 km scale combination as a reference, we estimate that the local sea level in [21], which used 2% and 50 km scales, may have been underestimated by about 0.05 m.

It has been suggested that our analysis could be used to investigate the impact of ice conditions in the estimation of the local sea level by looking at the variations between tracks. When we separated our analysis by tracks, however, we were not able to find consistent scale selections based on the two statistics used (R^2 and RMSE). We believe that this is related to the varying amount of data among the tracks. Nevertheless, using the largest track in our dataset (track 5), we showed that the lowest elevation method does not work very well when the ice is highly variable in space, resulting in an overestimation of the local sea level where the local sea level is low (outer part of the fluxgate >400 km) and an underestimation where the local sea level is high (inner part of the fluxgate < 400 km).

Our results provide the reference to derive the local sea level using the lowest elevation method especially for those altimetry data such as ICESat without simultaneous high-resolution imagery data. More research about the influence of using a reflectivity threshold is needed.

References

1. Holland, M.M.; Bitz, C.M. Polar amplification of climate change in coupled models. *Clim. Dyn.* **2003**, *21*, 221–232. [CrossRef]

2. Döscher, R.; Vihma, T.; Maksimovich, E. Recent advances in understanding the Arctic climate system state and change from a sea ice perspective: A review. *Atmos. Chem. Phys.* **2014**, *14*, 13571–13600. [CrossRef]

3. Meier, W.N.; Thomas, D. Losing Arctic sea ice: Observations of the recent decline and the long-term context. *Sea Ice* **2017**, Chap. 11. 290–303.

4. Simmonds, I. Comparing and contrasting the behaviour of Arctic and Antarctic sea ice over the 35 year period 1979–2013. *Ann. Glaciol.* **2015**, *56*, 18–28. [CrossRef]

5. Parkinson, C.; Cavalieri, D. Antarctic sea ice variability and trends, 1979–2010. *Cryosphere* **2012**, *6*, 871–880. [CrossRef]

6. Yi, D.; Zwally, H.J.; Robbins, J.W. ICESat observations of seasonal and interannual variations of sea-ice freeboard and estimated thickness in the Weddell Sea, Antarctica (2003–2009). *Ann. Glaciol.* **2011**, *52*, 43–51. [CrossRef]

7. Xie, H.; Ackley, S.; Yi, D.; Zwally, H.; Wagner, P.; Weissling, B.; Lewis, M.; Ye, K. Sea-ice thickness distribution of the Bellingshausen Sea from surface measurements and ICESat altimetry. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **2011**, *58*, 1039–1051. [CrossRef]

8. Zwally, H.; Brenner, A.; Farrell, S.; Laxon, S.; Yi, D. Deriving sea-ice freeboard height distributions and estimates of ice thickness from ICESat/GLAS laser altimetry. *AGU Fall Meet. Abstr.* **2003**, *2003*, C32A–0442.

9. Laxon, S.; Ridout, A.; Giles, K.; Willatt, R.; Wingham, D.; Hendricks, S.; Haas, C.; Beckers, J. Sea ice thickness from CryoSat2. *AGU Fall Meet. Abstr.* **2011**, *2011*, C41A-0374.

10. Giles, K.A.; Laxon, S.W.; Worby, A.P. Antarctic sea ice elevation from satellite radar altimetry. *Geophys. Res. Lett.* **2008**, *35*. [CrossRef]

11. Kurtz, N.T.; Markus, T. Satellite observations of Antarctic sea ice thickness and volume. *J. Geophys. Res.Oceans* **2012**, *117*. [CrossRef]

12. Zwally, H.J.; Yi, D.; Kwok, R.; Zhao, Y. ICESat measurements of sea ice freeboard and estimates of sea ice thickness in the Weddell Sea. *J. Geophys. Res.Oceans* **2008**, *113*. [CrossRef]

13. Farrell, S.L.; Laxon, S.W.; McAdoo, D.C.; Yi, D.; Zwally, H. Five years of Arctic sea ice freeboard measurements from the Ice, Cloud and land Elevation Satellite. *J. Geophys. Res.Oceans* **2009**, *114*. [CrossRef]

14. Price, D.; Rack, W.; Haas, C.; Langhorne, P.J.; Marsh, O. Sea ice freeboard in McMurdo Sound, Antarctica, derived by surface-validated ICESat laser altimeter data. *J. Geophys. Res.Oceans* **2013**, *118*, 3634–3650. [CrossRef]

15. Kurtz, N.T.; Farrell, S.L.; Studinger, M.; Galin, N.; Harbeck, J.P.; Lindsay, R.; Onana, V.D.; Panzer, B.; Sonntag, J.G. Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data. *Cryosphere* **2013**, *7*, 1035–1056. [CrossRef]

16. Wang, X.; Xie, H.; Ke, Y.; Ackley, S.F.; Liu, L. A method to automatically determine sea level for referencing snow freeboards and computing sea ice thicknesses from NASA IceBridge airborne LIDAR. *Remote Sens. Environ.* **2013**, *131*, 160–172. [CrossRef]

17. Krabill, W. *IceBridge ATM L1B Elevation and Return Strength, Version 2*; NASA DAAC at the National Snow and Ice Data Center: Boulder, CO, USA, 2013.

18. Krabill, W. *IceBridge ATM L2 Icessn Elevation, Slope, and Roughness, Data for 1993–2014*; NASA DAAC at the National Snow and Ice Data Center: Boulder, CO, USA, 2014.

19.   Dominguez, R. *IceBridge DMS L1B Geolocated and Orthorectified Images (IODMS1B, 2013)*; NASA DAAC at the National Snow and Ice Data Center: Boulder, CO, USA, 2010.

20.   Andersen, O.; Stenseng, L.; Piccioni, G.; Knudsen, P. The DTU15 MSS (mean sea surface) and DTU15LAT (lowest astronomical tide) reference surface. In Proceedings of the ESA Living Planet Symposium, Prague, Czech Republic, 9–13 May 2016.

21.   Tian, L.; Xie, H.; Ackley, S.F.; Tang, J.; Mestas-Nuñez, A.M.; Wang, X. Sea-ice freeboard and thickness in the Ross Sea from airborne (IceBridge 2013) and satellite (ICESat 2003–2008) observations. *Ann. Glaciol.* **2019**, 1–16. [CrossRef]

22.   Wang, X.; Guan, F.; Liu, J.; Xie, H.; Ackley, S. An improved approach of total freeboard retrieval with IceBridge Airborne Topographic Mapper (ATM) elevation and Digital Mapping System (DMS) images. *Remote Sens. Environ.* **2016**, *184*, 582–594. [CrossRef]

23.   Kern, S.; Spreen, G. Uncertainties in Antarctic sea-ice thickness retrieval from ICESat. *Ann. Glaciol.* **2015**, *56*, 107–119. [CrossRef]

24.   Kwok, R.; Cunningham, G.F.; Manizade, S.; Krabill, W. Arctic sea ice freeboard from IceBridge acquisitions in 2009: Estimates and comparisons with ICESat. *J. Geophys. Res. Oceans* **2012**, *117*. [CrossRef]

25.   Markus, T.T.; Neumann, A.; Martino, W.; Abdalati, K.; Brunt, B.; Csatho, S.; Farrell, H.; Fricker, A.; Gardner, D.; Harding, D.; et al. The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation. *Remote Sens. Environ.* **2017**, *190*, 260–273. [CrossRef]