Validation of CryoSat-2 SARIn Data over Austfonna Ice Cap Using Airborne Laser Scanner Measurements

Louise Sandberg Sørensen 1,†, Sebastian B. Simonsen 1,‡, Kirsty Langley 2, Laurence Gray 3, Veit Helm 1, Johan Nilsson 5, Lars Stenseng 1, Henriette Skourup 1,* and Malcolm W. J. Davidson 6

1 Department of Geodynamics, Technical University of Denmark, DTU-Space, 2800 Kgs. Lyngby, Denmark; slss@space.dtu.dk (L.S.S.); stenseng@space.dtu.dk (L.S.); hsk@space.dtu.dk (H.S.); rf@space.dtu.dk (R.F.)
2 ASIAQ-Greenland Survey, 3900 Nuuk, Greenland; kal@asiaq.gl
3 Department of Geography, Environment and Geomatics, University of Ottawa, Ottawa, ON K1N 6N5, Canada; laurence.gray@uottawa.ca
4 Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Department of Glaciology, 27568 Bremerhaven, Germany; veit.helm@awi.de
5 NASA Jet Propulsion Laboratory, 329C—Sea Level and Ice, California Institute of Technology, Pasadena, CA 91109-8099, USA; Johan.Nilsson@jpl.nasa.gov
6 European Space Agency, Mission Science Division, 2200 AG Noordwijk, The Netherlands; Malcolm.Davidson@esa.int
* Correspondence: ssim@space.dtu.dk; Tel.: +45-45259774
† These authors contributed equally to this work.

Received: 19 July 2018; Accepted: 22 August 2018; Published: 25 August 2018

Abstract: The study presented here is focused on the assessment of surface elevations derived from CryoSat-2 SARIn level 1b data over the Austfonna ice cap, Svalbard, in 2016. The processing chain that must be applied to the CryoSat-2 waveforms to derive heights is non-trivial, and consists of multiple steps, all requiring subjective choices of methods such as the choice of retracker, geo-relocation, and outlier rejection. Here, we compare six CryoSat-2 level-2 type data sets of surface elevations derived using different SARIn processing chains. These data sets are validated against surface elevation data collected from an airborne laser scanner, during a dedicated CryoSat validation experiment field campaign carried out in April 2016. The flight pattern of the airborne campaign was designed so that elevations were measured in a grid pattern rather than along single lines, as has previously been the standard procedure. The flight grid pattern was chosen to optimize the comparison with the CryoSat-2 SARIn elevation data, the location of which can deviate from nadir by several kilometers due to topography within the satellite footprint. The processing chains behind the six data sets include different outlier/error rejection approaches, and do not produce the same number of data points in our region of interest. To make a consistent analysis, we provide statistics from the validation of both the full data sets from each processing chain, and on only those data that all the six data sets provide a geo-located elevation estimate for. We find that the CryoSat-2 data sets that agree best with the validation data are those derived from dedicated land ice processing schemes. This study may serve as a benchmark for future CryoSat-2 retracker developments, and the evaluation software and data set are made publicly available.

Keywords: CryoSat-2; validation; Austfonna; airborne measurements; CryoVEx

1. Introduction

The CryoSat-2 (CS2) ku-band radar altimeter mission, launched in 2010, is currently providing unique and detailed information on the Earth’s ice-covered regions [1]. CS2 data have for instance
been used to assess elevation changes of Greenland [2–5], Antarctica [2,6], and smaller ice caps such as those in Iceland [7] and the Canadian Arctic [8]. Furthermore, CS2 data have also been used to study sea ice thicknesses [9–12], in-land waters [13–15], and sea-surface height and dynamic topography [16].

One current source of uncertainty in the CS2 elevation data products comes from differences in the processing chains applied to the Level-1b data to derive the CS2 Level-2 (L2) products which contain geolocated and time-tagged surface elevations. These typically include, for example, different waveform retrackers. Currently, several processing chains have been developed and applied by different research groups to study changes in land ice (e.g., [2–4,8]). In this study, six different CS2 L2 elevation data sets generated from different processing chains will be evaluated in terms of their agreement with an airborne laser scanner (ALS) data set used as validation. The validation data were collected over the Austfonna ice cap in the Svalbard archipelago located within one of the CS2 SARIn mode regions. Austfonna has been a test site for several dedicated CryoVEx (Cryosat Validation Experiment) airborne and in-situ field campaigns since 2003. During these campaigns, many types of airborne and in-situ data have been collected, including elevations from laser and radar, snow densities, and surface mass balance measurements [17].

In 2015–2016, the ESA funded the CryoVal Land-Ice project in which a thorough validation exercise was carried out and CS2 data were compared to available airborne data collected over land ice. One of the test regions was Austfonna, but it was found that the available airborne data did not form a basis for a solid statistical comparison. The reason was that in previous campaigns, only single lines following CS2 tracks were flown over Austfonna, yielding only few crossing points for elevation comparisons. The lack of crossing points is due to the nature of the CS2 SARIn measurements, and the fact that the locations of the retracked elevations followed the topography rather than the actual nadir position of the satellite. Therefore, in 2016 a new approach was tested for the CryoVEx campaign over Austfonna. Here, a dense grid aligned with the CS2 tracks was flown and ice heights mapped with ALS in order to provide a more suitable validation data set [18]. Here, we use this new airborne data set to assess the performance of six different CS2 elevation data sets. Studies like the present one can help evaluate the performance of the available SARIn processing chains for CS2 SARIn data, and their associated surface elevation uncertainty.

2. Study Area

Austfonna is the largest ice cap in the Svalbard archipelago with an area of 7800 km². It is a polythermal ice cap centered at 79°42′N, 24°E [19]. It has a dome-shaped topography, rising from sea-level to approximately 800 m a.s.l., and consists of well-defined drainage basins [20]. The northwest basins are land terminating, while the southeast basins calve into the Barents Sea. There is a strong precipitation gradient in most years, with snow fall typically ranging from 0.5 to 1 m in the west to 1.5–3 m in the east [21]. Mean bulk snow density, taken as an average of all snow pits from 2004 to 2013 is approximately \( 396 \pm 27 \text{ kg m}^{-3} \). Despite mean annual temperatures of \(-8.3^\circ\text{C}\), large temperature variations occur throughout the year and it is not uncommon for temperatures above 0 \( ^\circ\text{C}\) and rain events in winter [22]. In addition, summer melt occurs over the entire range of elevations, including the summit area.

3. CryoVEx ALS Data

In April 2016, as part of a CryoVEx airborne campaign, parts of the Austfonna ice cap were measured by near-infrared (904 nm) ALS. The measurements were carried out on 15–16 April 2016 from a Twin Otter airplane, which also carried the Airborne SAR/Interferometric Altimeter System (ASIRAS) radar [23]. The flight pattern and the measured surface elevations are shown in Figure 1. The waypoints for the flight were designed to be centered around two CS2 orbits (31,971 on 19 April and 32,066 on 26 April), and the extent of the grid was chosen so that it would cover most of the CS2 data from these orbits as well as part of a couple of other CS2 tracks crossing the ice cap in April. As seen in Figure 1, the grid for the western part (center line being track 32,066) was designed with closer line
spacing than the eastern grid (center line being track 31,917). This was to accommodate the different
topography in the two areas, which results in a greater deviation of the CS2 ground measurements
from the satellite nadir position in the eastern compared to the western part. The western grid has
a line spacing of 1 km while the eastern grid has 2 km line spacing. The easterly grid was partially
re-surveyed during the 2017 CryoVEx airborne campaign. For this validation analysis, the 2016
CryoVEx ALS data were averaged from their original swath width of 300 m and 1 × 1 m resolution to a
lower resolution of approximately 100 m × 100 m. This was done with a simple procedure, by deriving
the mean elevation in segments of a chosen size, with no outlier rejection applied. This averaging was
done to remove small-scale topography and to produce a spatial resolution comparable to CS2. The
data shown in Figure 1 are the averaged heights.

![Figure 1. The CryoVEx (Cryosat Validation Experiment) airborne laser scanner (ALS) height data
collected over Austfonna in April 2016. The elevations are given in meters above the WGS84 ellipsoid.
The black polygon outlines the main area (area 1) in which the validation is performed. In addition, this
area was split in two to provide additional information on how surface roughness interplays with the
inter-comparison of data sets. In the following, the area marked by the cyan polygon will be referred to
as area 2, and the southeast area will be referred to as area 3.](image)

4. CryoSat-2 L1b to L2 Processing

Here, we evaluate six different CS2 L2 data sets, referred to as ESA, JPL, AWI, AWI (DEM),
LARS NPP50, and UoO, against airborne ALS data over Austfonna. The following describe the details
of the main processing steps used in each of the six processing chains, to estimate surface elevations
over glaciated terrain from the CryoSat-2 L1b SARIn product.

To minimize any potential elevation differences between the satellite and airborne data
(e.g., snow fall), only L2 satellite data for the month of April 2016 were used.
4.1. ESA Baseline C L2i

The ESA CryoSat-2 processing facilities provide CS2 data products of different levels of complexity. In addition to the main Level-2 (L2) geophysical data record (GDR), a second L2 product is also provided in the form of the “In-depth”-product (L2i). This L2i product provides more parameters and flags than given in the traditional L2 GDR product. Here, we utilize the present Baseline C release of the ESA L2i data. In this processing chain, the SARIn area is retracked using the Wingham/Wallis model fit retracker [24]. The off-nadir position of the retracked point is (in SARIn areas) inferred using the across-track look angle, without the use of an a-priori slope model, as is the case for conventional radar altimetry (LRM). Further details of the data can be found in Systems [24].

4.2. JPL Baseline C L2

An independent processing chain of the CryoSat-2 L1b SARIn product was developed first at the Technical University of Denmark and later at the NASA Jet Propulsion Lab (JPL) by Johan Nilsson. In the JPL processing chain, each individual radar waveform is initially smoothed, using a zero-phase low-pass filter, to reduce the speckle noise in the waveform. Coherence values larger than one are set to zero and the coherence array (coherence versus range) is filtered using a $5 \times 5$ Wiener Filter. The interferogram is recreated and the real and imaginary parts are filtered individually using a wavelet de-noising approach. The differential phase is then recovered from the filtered real and imaginary parts of the interferogram, and is then unwrapped from the center of gravity of the radar waveform in each direction and merged back together. The radar waveform is retracked by finding the inflection point on the leading edge of the first surface return in the waveform. Some choices are made in the retracking to ensure good data quality:

- Only waveforms with an integrated power larger than $-170 \text{ dB}$ are used.
- The waveform is over-smoothed to remove small topographic signals.
- The first return is identified on the smoothed waveform and used to extract the low-pass-filtered leading edge. Peaks need to be two times larger than the waveform mean to qualify.
- Only first peaks inside the interval of gates 300–800 and waveforms with SNR $>0.5 \text{ dB}$ are used.
- The leading edge is over-sampled by a factor of 100 using linear interpolation.
- The inflection point is found by locating the maximum gradient of the leading edge.
- The range value of the extracted retracking point is used to find the corresponding phase and coherence values. Any measurement is rejected if the coherence value at the retracking point is less than 0.7.

The look angle is computed and corrected for the roll-angles error using the Gray et al. [25] roll-angle correction. The calculated look angle in combination with Earth geometry is used to calculate the echo return location on the ground. Phase ambiguities are identified and corrected for by comparing to an external digital elevation model (DEM) [19]. Here, a threshold value of 100 m is used for detection and if the difference is larger than this value $2\pi$ is added or subtracted from the phase and the return echo is relocated. Residual phase ambiguities are detected by comparing the along-track phase with a smoothed version. If a deviation larger than $1.5\pi$ is detected, the point is classified as an ambiguity and corrected. A more detailed description of the processor can be found in Nilsson et al. [3].

4.3. Alfred Wegner Institute (AWI) and AWI (DEM) Baseline C L2

Veit Helm at the Alfred Wegner institute (AWI) has developed an independent processing chain of the CryoSat-2 L1b SARIn product. Two different SARIn data sets from the processor were made available to this study. These are referred to as AWI and AWI (DEM). The processing chains of these are similar, and follow the one described in detail in Helm et al. [2]. This SARIn processing chain implies a threshold first-maximum retracker algorithm (TFMRA) and a relocation of the waveform position.
based on the coherence and phase difference, which are parameters included in the level 1B product. After the interferometric processing, the range is corrected for delays caused by the atmospheric refraction and tides. Following the TFMRA retracker approach, the AWI processing chain includes:

- Normalization of the waveform to its maximum.
- Calculation of the thermal noise and rejection of waveforms with a noise larger than 0.15.
- Over-sampling of the waveform by a factor of 10.
- Calculation of the smoothed waveform (*P*) with a boxcar average of width 15.
- Calculation of the first derivative *dP* using a three-point Lagrangian interpolation.
- Determination of the first local maximum and the location of the first gate that exceeds the threshold level at the leading edge of the first local maximum (chosen here to be 0.4 for SARIn).
- Determination of the exact leading edge position by interpolation between adjacent oversampled bins to the threshold crossing.

The further interferometric processing includes:

- Smoothing the phase in the complex domain and unwrapping in two directions.
- Defining the subset of range samples which cover the beam width.
- Rejection of data if the coherence value is less than 0.7.
- Application of mis-pointing-angle correction to baseline vector using data sets provided by ESA based on the study of Scaglia et al. [26]. This correction was only applied to the AWI (DEM) data set.

The slope-corrected geo-referenced point of closest approach (POCA) position is determined using the orbital position, the baseline vector, the range, and phase at the re-tracked position. The AWI (DEM) approach uses the same processing, but includes the use of an external reference DEM [19]. Here, we estimated two additional sets of geo-located POCA points by adding/subtracting 2π to/from the phase. The POCA position with the lowest offset to the reference DEM is chosen. Further along-track processing as suggested by Nilsson et al. [3] was not applied.

### 4.4. University of Ottawa (UoO) Baseline C L2

The final dedicated land ice processing chain available for this study was provided by Laurence Gray at the University of Ottawa (UoO). The UoO processing chain starts with initial quality control of the CryoSat-2 L1b SARIn, which includes checking the magnitude of the average of the first few values in each waveform. If the dB value of this average is greater than −150, then the waveform is ignored as it is likely that the altimeter tracking loop has missed the surface. The waveform power and differential phase provided in the ESA L1b SARIn file are used to create in-phase and quadrature components which are then low-pass-filtered and used to create a smoothed version of the phase and waveform. While this stage increases the cross-track footprint size, the lower phase noise improves the estimation of the cross-track look angle and therefore the geolocation of the POCA. The retracker algorithm looks for the first significant “leading edge” in the waveform, up-samples this region by a factor of 10, and uses the maximum slope position as the position of the POCA. This approach, using the first significant leading edge of the waveform, is similar to that used in the AWI and JPL processors. Initially, the phase at the POCA is used to calculate the look direction with respect to the line connecting the centers of the two receive antennas (the interferometric “baseline”) using the pre-launch measurement of the baseline [25]. Satellite state vectors, timing data, and the orientation of the baseline in the “CryoSat reference frame” [27] are provided in the L1b file, and these data are used to geo-locate the POCA. Solutions are derived for the phase at the estimated POCA position in the waveform, and for this phase +2π and −2π. Comparison with the height of a reference digital elevation model [19] is used to select the most likely of the three solutions. Further editing can be done based on some of the data saved after the processing. For this test, solutions were rejected if the coherence at the POCA point was less than 0.7, if the ratio of the maximum waveform power to
the average of the first five values was too low (<6 dB), or if there was no clear leading edge in the waveform. Further details of the UoO SARIn data processing have been described in Gray et al. [8] and Gray et al. [25].

4.5. LARS Baseline C L2

Besides the dedicated land ice processing chains described in the previous sections, we also include a L2 data set generated from a processing chain developed for ocean applications. This data set was generated at the Technical University of Denmark (DTU) using the LARS Advanced Retracking System (LARS) which was originally constructed to retrack CryoSat-2 waveforms in the Arctic Ocean in the presence of sea ice using the empirical Offset Centre of Gravity (OCOG), Threshold, and Beta-5 retrackers. The outputs from the system were tested against shipborne gravity data and showed promising results for CryoSat-2 and the new SAR mode [28]. Since then, additional retrackers have been added to LARS, which now consists of more than 10 empirical and physical retrackers primarily targeted at open ocean [29–31], coastal regions [32], rivers [14], and lakes [13]. For this study, the Narrow Primary Peak Retracker with 50% threshold (NPP50) [31,33] was modified following Villadsen et al. [14] to include geo-location following Abulaitijiang et al. [32], Armitage et al. [34]. Key-points in the processing chains includes:

- Thermal noise is calculated and subtracted. Negative values are truncated to zero.
- The start and stop thresholds are determined as the standard deviation of the power in alternating and consecutive bins, respectively.
- The primary peak is here defined at the first peak above 25% of the maximum power. All retracking is then preformed on this subwaveform.

5. Validation Method

The CS2 heights \((H_{CS2})\) derived from each processing chain were validated using ALS data collected during the dedicated CryoVEx validation campaign on 15 and 16 April 2016 (Section 3) using the following method.

For each CS2 L2 data observation location, the corresponding validation elevation value is found by means of simple linear interpolation of the elevations measured by ALS \((H_{ALS})\) and the difference between the two are calculated as: \(\Delta H = H_{CS2} - H_{ALS}\). Using this method, no maximum distance to the actual ALS measurements is required, but due to the nature of the grid flown, this distance will be at most 1 km. In order to base the statistics of the \(\Delta H\) estimates only on interpolated values of the ALS data and avoid errors introduced by extrapolation, we only estimate elevation differences within a polygon closely outlining the area covered by the ALS measurement grids, and therefore exclude most of the single east–west-oriented line (Figure 1). For each of the six CS2 L2 data sets, \(\Delta H\) are evaluated, and a mean, median, and standard deviation based on the histograms are derived. This is done for all available CS2 L2 data, and for the subset of points where all six data sets provide an elevation estimate. This common data set was found from a comparison of the nadir locations of the data points, since their geo-locations will not necessarily be identical due to the differences in processing strategies.

6. Results

Elevation differences \((\Delta H)\) between the six CS2 L2 data sets and the surface elevations mapped with ALS are shown in Figure 2. The CS2 data locations are indicated by black dots, and the polygon that marks the full interpolation area for the ALS data (Area 1) is shown by a black line.

The histograms of the \(\Delta H\) estimates for all 600 common points for the six data sets are shown in Figure 3. The range \([-50 \text{ m}, 50 \text{ m}]\) of the \(\Delta H\) values in the histogram are given in bins of 1 m (shown in red), while for values outside this range, the bin size is 25 m (shown in blue). Only UoO did not have absolute differences larger than 50 m, whereas the others had 7, 16, 17, 6, and 40 for AWI, AWI (DEM), ESA, JPL, and LARS respectively.
Figure 2. Elevation differences ($\Delta H$) between the six CryoSat-2 Level-2 (CS2 L2) data sets and the surface elevations mapped with ALS over Austfonna, April 2016. Black polygons indicate the area within which the $\Delta H$ comparison was made. The CS2 point of closest approach (POCA) data locations are marked with small black dots outside the polygon that marks the area for interpolation to ALS data is shown in green. AWI: Alfred Wegner Institute; DEM: digital elevation model; JPL: NASA Jet Propulsion Lab; LARS NPP50: LARS Advanced Retracking System Narrow Primary Peak Retracker with 50% threshold; UoO: University of Ottawa.

The statistics of the derived $\Delta H$ values from all six CS2 L2 data sets, described by their mean, median, and standard deviation, are given Table 1. The statistics are based on all available $\Delta H$ values for a given data set, and it is seen that the number of $\Delta H$ values differed, which is a consequence of the different outlier detection strategies implemented in the individual processing chains. The AWI data set provided the most $\Delta H$ values (828), while the JPL provided the fewest (725). The table also contains the statistics for only those $\Delta H$ data points that had common nadir points in all six data sets (the numbers in parenthesis). Therefore, the main results shown in Table 1 give the performance of each full data set, while the numbers in parentheses show the statistics which are best suited for a direct comparison of the performance of the processors. The absolute lowest values of mean, median, and standard deviation are highlighted in bold. From the full data set analysis (Table 1) we see that the ESA data set provided a low mean value, but the largest standard deviation. We see from the statistics of the common data points (Table 1) that the heights derived from the JPL SARIn processor generally agreed best with the ALS data, but that also the AWI and UoO processors provided results...
with good validation statistics. The ESA, JPL, and UoO data were the only data sets that yielded a negative median of $\Delta H$, indicating that the CS2 heights were generally lower than the ALS heights.

**Figure 3.** Histograms of elevation differences ($\Delta H$) between the six CS2 L2 data sets and the surface elevations mapped with ALS over Austfonna, April 2016. These histograms show the 601 $\Delta H$ that all eight CS2 L2 data sets have in common. The red part of the histograms is provided in 1 m bins while the blue is in 25 m bins.

**Table 1.** $\Delta H$ statistics for the six different CS2 data sets. Additionally, the statistics for data points in common for all data sets are provided in parenthesis. The areas refer to Figure 1.

<table>
<thead>
<tr>
<th>Area</th>
<th>CS2 Data Set</th>
<th>ESA</th>
<th>JPL</th>
<th>AWI</th>
<th>AWI (DEM)</th>
<th>LARS</th>
<th>UoO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td># of $\Delta H$</td>
<td>777 (600)</td>
<td>725 (600)</td>
<td>787 (600)</td>
<td>828 (600)</td>
<td>768 (600)</td>
<td>752 (600)</td>
</tr>
<tr>
<td></td>
<td>Mean (m)</td>
<td>2.80 (3.07)</td>
<td>1.14 (0.07)</td>
<td>4.65 (3.62)</td>
<td>4.42 (4.55)</td>
<td>13.64 (14.18)</td>
<td>0.93 (0.61)</td>
</tr>
<tr>
<td></td>
<td>Median (m)</td>
<td>−1.11 (−1.12)</td>
<td>−0.28 (−0.19)</td>
<td>2.04 (2.01)</td>
<td>2.34 (2.33)</td>
<td>5.53 (5.62)</td>
<td>−0.31 (−0.35)</td>
</tr>
<tr>
<td></td>
<td>Std. dev. (m)</td>
<td>30.28 (30.71)</td>
<td>11.71 (6.65)</td>
<td>11.84 (6.27)</td>
<td>18.45 (16.84)</td>
<td>43.52 (45.22)</td>
<td>4.80 (4.38)</td>
</tr>
<tr>
<td>2</td>
<td># of $\Delta H$</td>
<td>509 (417)</td>
<td>470 (417)</td>
<td>509 (417)</td>
<td>512 (417)</td>
<td>494 (417)</td>
<td>497 (417)</td>
</tr>
<tr>
<td></td>
<td>Mean (m)</td>
<td>−0.76 (−1.33)</td>
<td>−0.48 (−0.33)</td>
<td>4.31 (2.02)</td>
<td>2.72 (2.35)</td>
<td>4.89 (4.90)</td>
<td>−0.56 (−0.56)</td>
</tr>
<tr>
<td></td>
<td>Median (m)</td>
<td>−1.04 (−1.01)</td>
<td>−0.34 (−0.27)</td>
<td>1.63 (1.58)</td>
<td>2.04 (2.06)</td>
<td>5.53 (5.57)</td>
<td>−0.87 (−0.82)</td>
</tr>
<tr>
<td></td>
<td>Std. dev. (m)</td>
<td>14.63 (3.15)</td>
<td>2.93 (1.87)</td>
<td>12.57 (1.94)</td>
<td>6.61 (1.94)</td>
<td>19.19 (19.52)</td>
<td>1.97 (1.87)</td>
</tr>
<tr>
<td>3</td>
<td># of $\Delta H$</td>
<td>268 (169)</td>
<td>258 (169)</td>
<td>278 (169)</td>
<td>318 (169)</td>
<td>274 (169)</td>
<td>256 (169)</td>
</tr>
<tr>
<td></td>
<td>Mean (m)</td>
<td>9.57 (14.28)</td>
<td>4.00 (1.01)</td>
<td>5.27 (6.41)</td>
<td>7.15 (6.83)</td>
<td>29.43 (37.33)</td>
<td>3.84 (3.45)</td>
</tr>
<tr>
<td></td>
<td>Median (m)</td>
<td>−1.43 (−1.89)</td>
<td>−0.01 (−0.01)</td>
<td>3.78 (3.93)</td>
<td>3.99 (4.38)</td>
<td>5.51 (6.46)</td>
<td>1.54 (1.19)</td>
</tr>
<tr>
<td></td>
<td>Std. dev. (m)</td>
<td>46.72 (56.10)</td>
<td>18.91 (6.07)</td>
<td>10.33 (6.46)</td>
<td>28.35 (6.63)</td>
<td>65.25 (74.08)</td>
<td>6.88 (6.87)</td>
</tr>
</tbody>
</table>

7. Discussion

From the spatial maps of the CS2-ALS elevation differences (Figure 2), we can see more specifically where the different data sets’ performances are good and where they are degraded. In general, we see that all six data sets showed better agreement with the western-most ALS grid (area 2), compared to the eastern one (area 3). This is probably due to the more gentle topography in the Western region compared to the Eastern part. The latter covers a surging glacier [35], and the surface here is characterized by rough topography with many deep crevasses, reflected in the standard deviation of area 3 (Table 1). This rough terrain might be challenging for the processors, and in addition may not sample the topography accurately enough with a 2 km grid spacing. To avoid interpolating the ALS data, we attempted to derive statistics on only overlapping or closely separated ALS and CS2 data, but this resulted in a very small data set, which was not suitable for a robust validation.

Due to the nature of the CS2 SARIn measurements, the precise geolocation of the measurements is a very important part of a given processing chain. It is clear by visual inspection of the data that this procedure differed substantially between the processors considered here. Figure 4 shows the large
variability in the geolocations of CS2 measurements from the different L2 data sets. The figure shows both the locations of the derived heights (colors indicating different processors) as well as their nadir position (black). In the inset in the figure, we show for clarity the data locations for a single CS2 track. Here it is seen that along most of the track, the locations of the different data sets agreed well, but also that in specific regions individual processors seemed to struggle and produce locations that clearly deviated. For example in the southern end of the selected track, we see that the ESA echo locations deviated from the path of the other data sets. A correct geolocation is probably one of the most important parts of a successful SARIn processor, and this could be the focus of future studies. The AWI (DEM), JPL, and UoO processors had the best validation statistics, and common for these processors was a general good agreement of the echo-geolocation. Furthermore it is crucial to implement an appropriate retracker or threshold, as this alone has been shown to affect the performance over Greenland [3]. Common for these three processing chains is the use of a reference DEM to guide the interferometric processing in order to avoid large outliers caused by $2\pi$ phase wrapping. This was shown by the reduced standard deviation of AWI (DEM) compared to AWI for areas 1 and 3.

Figure 4. Data geolocations for the six CS2 L2 data sets, indicated here by different colors. Shown in black are their common nadir locations. The inset shows the geolocation differences caused by different processing strategies on a single CS2 track.

Our analysis shows that, with the exceptions of ESA, JPL, and UoO data, the mean and median of $\Delta H$ from the other processors (both complete and common data) were positive (Table 1). This finding is surprising since, due to possible penetration of the radar signal into the snow [36,37] we would expect these to be negative (i.e., CS2 lower than ALS). The ALS data were averaged to $100 \times 100$ m resolution. In a heavily crevassed area such as area 3, this would result in a lower average surface, since the elevation originating from within the crevasses is included. Conversely, the CS2 radar altimeter measures the highest point in the footprint. This was confirmed by the positive $\Delta H$ in area 3 and negative $\Delta H$ in area 2 mapped by UoO. However, the other CS2-data showed results which cannot be explained by the presence of crevasses. At present, we do not have a good explanation for this finding, although one possible explanation could be the presence of a bias in the CS2 L1b data itself.
The presence of a potential bias does not affect the conclusions drawn in this study, as it does not affect the standard deviations.

The criteria of using only CS2 data from the same month as the ALS data were collected should minimize any actual elevation differences between the two data sets caused by, for example, snow fall. However, such actual elevation changes cannot be eliminated completely. Any influence of actual surface elevation differences occurring between the time of the satellite and airborne measurement will affect the statistics, but since they are the same for all six data sets, the inter-comparison is still valid. Additionally, we investigated if a slope-dependent bias was present in the $\Delta H$ data but this was not the case.

It should be emphasized that the conclusions on the preferred processing scheme drawn in this study cannot be directly transferred to other locations, since climatology and snow characteristics might differ and this would likely have an impact on the performance of each processing strategy, as the radar altimetry data are heavily dependent on these parameters.

8. Conclusions

Data from six different CryoSat-2 SARIn processing chains (ESA, JPL, AWI DEM, AWI, LARS NPP50, UoO) were tested and compared through a comparison of their respective cross-over difference statistics with ALS data over the Austfonna ice cap collected in April 2016, during a dedicated ESA CryoSat-2 validation campaign. Our study shows that the heights derived from the AWI, JPL, and UoO SARIn processors agreed best with the ALS data. These processors were designed specifically for the land ice application, and all three identify and use analysis of the first significant leading edge, rather than using the entire waveform. We believe that the improved performance of these three processors is also related to the fact that the POCA is derived from only the leading edge and not the whole waveform. The LARS processors, which were designed for other purposes, provided the data sets which deviated most from the ALS validation data. For this reason, we found that dedicated land ice SARIn processing chains are needed to derive reliable height estimates. The major difference between the processing chains is in how data are geo-located, and this had a large impact on the validation statistics. The geo-location of the data is dependent on the retracker and phase filtering applied since these determine the look angle that dictates the across-track position of the measurement. Here, the preferred processing chains (AWI (DEM), UoO, and JPL) use a reference DEM to guide the interferometric processing to avoid large outliers caused by $2\pi$ phase wrapping.

As all presented data set and validation procedures are made available online, we see this study as a benchmark for future retracker development.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/10/9/1354/s1. All CS2 and ALS data are available in the supporting on-line material (SOM). The SOM also contains the Python code used to make the statistics, to allow future retracker developments to be directly benchmarked against this study.

Author Contributions: L.S.S. and S.B.S. conceived and designed the study based on the findings in the ESA CryoSat-2 Validation Land-ice project. S.B.S., V.H., J.N., L.S., and L.G. contributed with CryoSat-data; All authors contributed to writing the paper and discussed the results. H.S., L.S.S. and S.B.S. participated in ALS field campaign and processing; R.F. organized the overall field campaign logistics in 2016, along with M.W.J.D. as the acting ESA-officer.

Funding: This study was founded through ESA CryoSat land validation, contract 4000109790/13/NL/CT/ab. The German Ministry of Economics and Technology (grant 50EE1331 to V. Helm) has supported this work.

Acknowledgments: This study was founded through ESA CryoSat land validation, contract 4000109790/13/NL/CT/ab. The German Ministry of Economics and Technology (grant 50EE1331 to V. Helm) has supported this work. The mispointing-angle correction applied to to baseline vector was downloaded from ESA at https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/cryosat/news/-/article/updated-cryosat-dataset-with-corrected-mispointing-angle-for-aberration-of-light.

Conflicts of Interest: The authors declare no conflict of interest.
References


2. Helm, V.; Humbert, A.; Miller, H. Elevation and elevation change of Greenland and Antarctica derived from CryoSat-2. *Cryosphere* 2014, 8, 1539–1559. [CrossRef]


17. CSAG. *CryoSat Calibration & Validation Concept*; Technical Report November, ESA/CPOM; Centre for Polar Observation & Modelling Department of Space & Climate Physics, University College London: London, UK, 2001.


