A New Method for Iceberg Tracking Using Contour Matching

Lotte H. de Boer

Delft University of Technology

TUDelft

A New Method for Iceberg Tracking Using Contour Matching

by

Lotte H. de Boer

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Tuesday February 21, 2023 at 1:00 PM.

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TUDelft

Abstract

Icebergs drifting through the Southern Ocean release fresh water and nutrients. This has local impacts on surrounding ecosystems and sea ice formation. On a global scale, salinity patterns and ocean circulation are affected. In addition, studying icebergs as a proxy for ice shelves in a warming climate can help predict future climate impacts and sea level rise. Furthermore, drifting icebergs can pose a threat to ship navigation and offshore projects. In the past, icebergs have been tracked mostly manually, a time-consuming and labour-intensive task. The most widely used data source for this is Synthetic Aperture Radar (SAR), as icebergs often have a much higher backscatter than their surroundings. A few attempts have been made to automatically track icebergs, but these methods do not allow tracking of icebergs that are only partially visible in a satellite image. In this study, a new method is proposed based on partial contour recognition using the contours' curvature, a technique derived from the matching of ancient pottery fragments. Since the automatic tracking of multiple icebergs requires a large amount of data and computational resources, the web-based environment of Google Earth Engine is used.

The new method, called the Contour Curvature (CC) method, is based on three main steps. (1) Detection of icebergs using Simple Non-Iterative Clustering (SNIC) in combination with a threshold function. (2) The icebergs targets are filtered using an area and solidity filter. (3) Among the remaining targets, the best match is selected by comparing the curvature function of the contour with the reference iceberg. The performance of the algorithm is tested by automatically tracking 15 icebergs and comparing the results to the existing Centroid Distance Histogram (CDH) method. The overall performance of the CC method can be attributed in large part to the inclusion of the area and the solidity filter, with the latter serving as an overall shape filter. For small icebergs (< 10 km^2), both the CC and CDH method perform poorly, due to the abundance of icebergs in this range. For medium to large icebergs (10 to 1000 km^2), the methods show similar performance with one method occasionally outperforming the other method. For large icebergs (> 1000 km^2), the CC method performs better. Since these icebergs are often only partially visible, this leads to strong deviations in the histogram used in the CDH method, making this method less suitable for these situations. Since the CC method allows for partial contour recognition, these icebergs can still be identified. Furthermore, due to the wide variety of backscatter conditions, the detection method occasionally fails to distinguish icebergs from their surroundings.

Contents

Introduction

1.1. The Impact of Drifting Icebergs

Drifting icebergs move around Antarctica in the Southern Ocean, affecting the environment in a variety of ways. Icebergs are, next to basal melting, the primary source of incoming freshwater flux in the Southern Ocean (Depoorter et al., 2013; Silva et al., 2006). Cold fresh water from melting icebergs disrupts temperature and salinity patterns (Schodlok et al., 2006), affecting global ocean circulation (Barbat, Rackow, et al., 2019). Locally, this meltwater allows for increased sea ice production (Braakmann-Folgmann et al., 2022; Mazur et al., 2017; Pauling et al., 2016). Furthermore, grounded icebergs alter local ocean dynamics by acting as a barrier, disrupting circulation and the benthic ecosystem while also blocking passages for animals such as penguins (Braakmann-Folgmann et al., 2022; Schodlok et al., 2006). In addition, nutrients that have been trapped in the ice for thousands of years have an effect on biological activity. These nutrients have the potential to fertilise the upper layer of the ocean (Braakmann-Folgmann et al., 2022; Schodlok et al., 2006). Knowing the precise positions and trajectories of the icebergs, facilitates the ability to estimate fresh water input and study the effects on the ecosystem and ocean dynamics (J. S. Budge & Long, 2018; Schodlok et al., 2006).

Drifting icebergs can form a serious threat to offshore structures and ship navigation (Koo et al., 2021; Yulmetov et al., 2016; Zhan et al., 2018). Large tabular icebergs can range in length from a few kilometres up to several hundred kilometres (Braakmann-Folgmann et al., 2022; Stuart & Long, 2011). When these massive icebergs collide with offshore structures, severe damage can occur. If the locations of icebergs are known and their trajectories can be estimated, future collisions may be avoided.

The Antarctic Ice Sheet is one of the largest ice reservoirs on earth and therefore has the potential to have a large contribution to global sea level rise if the climate continues to warm (Mazur et al., 2017). The calving of icebergs at ice shelf margins or glacier tongues accounts for half of all ice loss from Antarctica (Braakmann-Folgmann et al., 2021). The future contribution of the Antarctic Ice Shelf to global sea level rise is still highly uncertain, but if average global temperatures rise with 5 degrees Celsius, the ice sheets could contribute up to 2 metres of sea level rise by 2100 (Bamber et al., 2019). More information on how ice shelves will react to a warmer climate is needed to make better predictions about the amount of sea level rise in the future. As large tabular icebergs are similar to ice shelves in their physical characteristics and stress environments (Scambos T, 2005), they are suitable for exemplar studies. Tabular icebergs that slowly drift to warmer areas can be used as a proxy for ice shelves that are gradually warming as a result of global warming (Braakmann-Folgmann et al., 2022).

1.2. Tracking of Antarctic Icebergs

In the past, icebergs have been tracked manually by attaching GPS-buoys to drifting icebergs (Schodlok et al., 2006). Since 1978, the US National Ice Center has used Synthetic Aperture Radar to track icebergs (J. S. Budge & Long, 2018; Center, n.d.). Manually inspecting and registering icebergs, is however, still a labour and time-consuming task. Successful methods have been developed for the automatic detection of icebergs in SAR data using image segmentation (Barbat, Wesche, et al., 2019;

Koo et al., 2021). However, automatic tracking of icebergs is not yet widely studied and only a few methods have been tested in recent years (Barbat et al., 2021; Koo et al., 2021).

The automatic tracking of icebergs is a difficult task because icebergs come in a variety of shapes and sizes and drift with ocean and wind currents (Koo et al., 2021). Icebergs are constantly shrinking and changing shape as a result of fracturing and gradual melting caused by warmer water temperatures, wave interactions and current and wind regimes (Barbat, Wesche, et al., 2019). These factors make it difficult to create a general algorithm that can be applied on the entire range of icebergs varieties.

Barbat et al., 2021 has achieved to automate the tracking of more than 400 icebergs in the Weddell Sea using SAR imagery. To re-detect icebergs in different time frames, one-dimensional signatures based on the distances from the iceberg's border to the centroid are created. This is done in a circular pattern, producing a vector with a length of 360 (one value per degree), which is then sorted to produce a rotation-invariant signature. Although this method has proven to be effective on a variety of icebergs ranging in size from 3.4 to 3612 km^2, one great disadvantage is that the algorithm is used locally, necessitating large storage and processing capacities.

The use of Google Earth Engine (GEE) is a technique for overcoming the need for large data storage and processing capacity (Koo et al., 2021). GEE is a cloud computing platform that allows for the access and processing of large amounts of data (Google Earth Engine, n.d.-b). Koo et al., 2021 demonstrated that GEE is a suitable tool for the tracking of iceberg B43 using the freely available Sentinel-1 SAR data. In the used method, iceberg signatures are created by forming a histogram based on the distance between each pixel in the iceberg and the the iceberg's centroid. The above-mentioned methods developed by Barbat et al., 2021 and Koo et al., 2021 have the disadvantage that the created signatures are dependent on the iceberg's centroid and thus only work for icebergs that are fully visible in SAR images. When an iceberg is only partially visible, the centroid is displaced, resulting in significant deviations in the signatures and, as a result, failure of the method.

Methods that are independent of the centroid position should be considered when tracking icebergs that are not always fully visible in a single satellite image. A partial shape recognition method is described by Liu and Srinath, 1990, where shapes are matched based on their contour. Before comparing different shapes, this method divides the shape contours into different segments. In this way, shapes can be recognised even if they are partially distorted. Matching the different shapes is done based on the curvature of their contour. The minima and maxima of the curvature function are then used to cut the contour into different segments. The problem has become rotation and translation independent by employing the curvature function. The method has been shown to be successful for various test shapes with a moderate amount of noise, round shapes and shapes with sharp edges (Liu & Srinath, 1990).

Da Gama Leitão and Stolfi, 2002 also describe using the curvature function for shape matching. This method focuses on the edge matching of two-dimensional flat or curved surfaces with irregular, weathered edges. The error between the curvature functions of different fragment edges is calculated, and this determines whether two edges are a true match. When viewed from above, tabular icebergs can be thought of as two-dimensional shapes with weathered edges. Their contours are constantly changing over time due to melt and weathering caused by wave interaction, and may appear also different in SAR images due to shadowing or melt ponds which alter the backscatter of the iceberg.

1.3. Research Objective

In this study, a new method is proposed and its performance is tested for icebergs of varying shapes and sizes under various backscatter conditions. To reduce the large amount of available icebergs, preliminary area and solidity filters are included. Furthermore, to compare the performance of the method, a benchmark study is performed with an already existing method. The goal of this research can be summarized by the following main objective:

> *To construct and test a method to detect and track icebergs, by matching their contour segments based on the curvature function.*

To test the performance and robustness of the method, the following research questions are addressed:

1. What is the influence of different iceberg properties and surroundings on the performance of the method?

2. How large is the contribution of the area and solidity filter on the ability to correctly select a match?

3. To what extent is the algorithm able to recognize icebergs that are only partially visible in a satellite image?

4. How does the performance of the method compare to an already existing method?

This study is structured as follows. Chapter 2 describes the used data set as well as the defined research area and the tracked icebergs. Chapter 3 describes the steps of the constructed method and an existing method. Followed by the procedure used to compare both methods. Chapter 4 presents the tracking results of both methods and compares their performance. Chapter 5 contains a discussion of the outcomes and examines the limitations and how to potentially overcome these. The findings of the study are summarized in chapter 6.

<div align="right">

2

</div>

<div align="right">

Data

</div>

Synthetical Aperture Radar (SAR) is a commonly used data source for detecting and tracking icebergs. For this study, Sentinel-1 SAR Ground Range Detected (GRD) is used, in the area of the Weddell Sea. The data is processed in the Google Earth Engine environment which has the chosen datasets readily available. The choices for this programming environment and the data sets, as well as the study area are explained below.

2.1. Programming Platform; GEE

The cloud programming platform Google Earth Engine (GEE) is used. One of the great advantages of GEE is that large amounts of data are freely available and can be manipulated in the cloud without the need of storing large datasets on a local computer (Google Earth Engine, n.d.-b). This saves a lot of time and computing resources (Koo et al., 2021). The default programming language inside the GEE code editor interface is JavaScript but the platform is easily integrated with python, allowing for greater functionality.

2.2. Sentinel-1 SAR GRD Images

An advantage of SAR is the ability to look through clouds, allowing the acquisition of data under almost all meteorological conditions (Mazur et al., 2017). Furthermore, it is an active operating system, meaning it is independent of sunlight conditions. In general, icebergs have a higher radar backscatter intensity than their surroundings, which consist mostly of sea ice and water (Mazur et al., 2017). This is because the relative dielectric constant of non-saline ice (of which icebergs consist) is low and the backscatter of an iceberg consists of surface as well as volume scattering (Mazur et al., 2017), making the iceberg appear bright in SAR images. Water and sea ice, on the other hand, have a much lower backscatter and thus appear darker in the SAR images. This makes it easier to distinguish the icebergs from their surroundings.

The Sentinel-1 SAR Ground Range Detected (GRD) is used, which is a collection consisting of multiple scenes. Each of the scenes contains multiple resolutions (10, 25 or 40 meters) and multiple polarisations (VV, HH, VV+VH and HH+HV) (Google Earth Engine, n.d.-a). As the tracked icebergs have lengths of multiple kilometres, all three resolutions are sufficient. For the polarisation, the HH-band is chosen, as it is available for most of the images (Koo et al., 2021). The SAR instrument operates during day and night at a wavelength of 5.5 cm (ESA, n.d.).

The main operation mode has a wide swath of 250 km (ESA, n.d.). The Sentinel-1 constellation, consisting of two satellites, is in a near-polar, sun-synchronous orbit with a repeat cycle of 6 days (ESA, n.d.). However, near the poles, orbits are overlapping which reduces the return period over Antarctica to once every two to three days in many regions (Koo et al., 2021). Sentinel-1 data is available from April 2014 and is updated daily, with a one day delay. Because there is only one day delay in the availability of the data, the dataset is suitable for near real-time tracking. Before the data is uploaded, it is preprocessed using the Sentinel-Toolbox which produces the Ground Range Detected product, which

is a calibrated, ortho-corrected product (Google Earth Engine, n.d.-a). The Sentinel-Toolbox preprocessing consists of three steps; thermal noise removal, radiometric calibration and terrain correction (Google Earth Engine, n.d.-a).

2.3. Area of Study

The Weddell Sea is located on the west side of Antarctica, the main ice shelves located in the Weddell Sea are the Larsen, Filchner-Ronne and the Eastern Weddell Ice Shelves (Timmermann, 2002), which are sources for the creation of icebergs. Other icebergs within the Weddell Sea originate from the East and are driven into the Weddell Sea by the Antarctic Coastal Current (Barbat, Wesche, et al., 2019; Vernet et al., 2019). The icebergs tracked in this research have trajectories in the region between 12°East and -60°West and between -60°and -76°South, this region is marked in blue in Figure 2.1.

The circulation in the Weddell Sea is dominated by the Weddell Gyre (Barbat, Wesche, et al., 2019) which is driven by wind forcing and is located south of 55-60°South and between 60°West and 60°East (Vernet et al., 2019). Iceberg trajectories in the Weddell sea mostly follow the Weddell Gyre but are also affected by wind regimes.
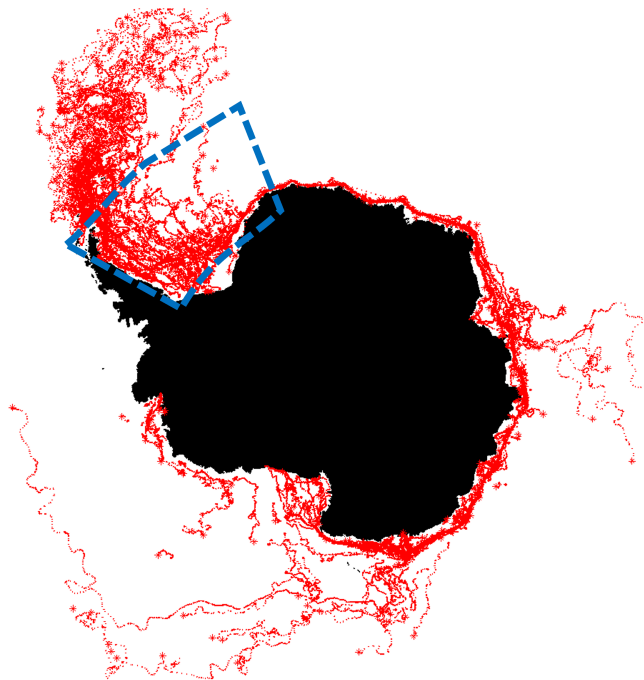

Figure 2.1: The Antarctic continent with iceberg trajectories (in red) documented by J. Budge et al., 2022. The extent of the search area containing the Weddell Sea is indicated in blue.

2.4. Tracked Icebergs

To test the performance of the method, 15 icebergs of different sizes and solidity are tracked for one year, between April 2015 and December 2018. A number of 15 icebergs is sufficient to cover a wide range of properties while keeping the manual and cloud data processing at a reasonable level. The areas of these icebergs range from 3 to 5600 km^2 and the solidity values range from 0.75 to 0.94. The solidity is a shape measure which is defined as the area of the iceberg divided by the convex hull, this principle is further explained in subsection 3.1.2. How the areas and solidity values of the different icebergs relate to each other is shown in Figure 2.2.

Further properties and the exact tracking start and end dates, as well as coordinates can be found in Appendix B.
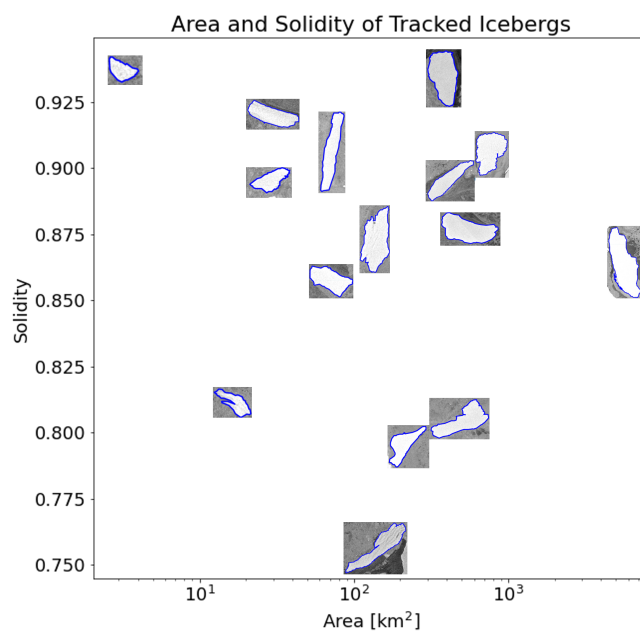
Figure 2.2: The area [km^2] and solidity of the 15 tracked icebergs viewed in a scatterplot.

3

Methodology

In this chapter, the Contour Curvature (CC) method as well as the Centroid Distance Histogram (CDH) method are described. The first is a newly constructed algorithm based on iceberg detection by Simple Non-Iterative Clustering (SNIC) in combination with a threshold function as described by Koo et al., 2021, followed by contour recognition based on principles of curvature matching as described by Liu and Srinath, 1990 and Da Gama Leitão and Stolfi, 2002. The CDH method is an existing method developed by Koo et al., 2021 using SNIC for iceberg detection, followed by a matching step using centroid distance histograms which describe the size and shape of each iceberg. Figure 3.1 depicts both methods schematically, with the CC method on the left (a) and the CDH method on the right (b). Steps that are the same for both methods are shown in blue, steps that differ are shown in orange.
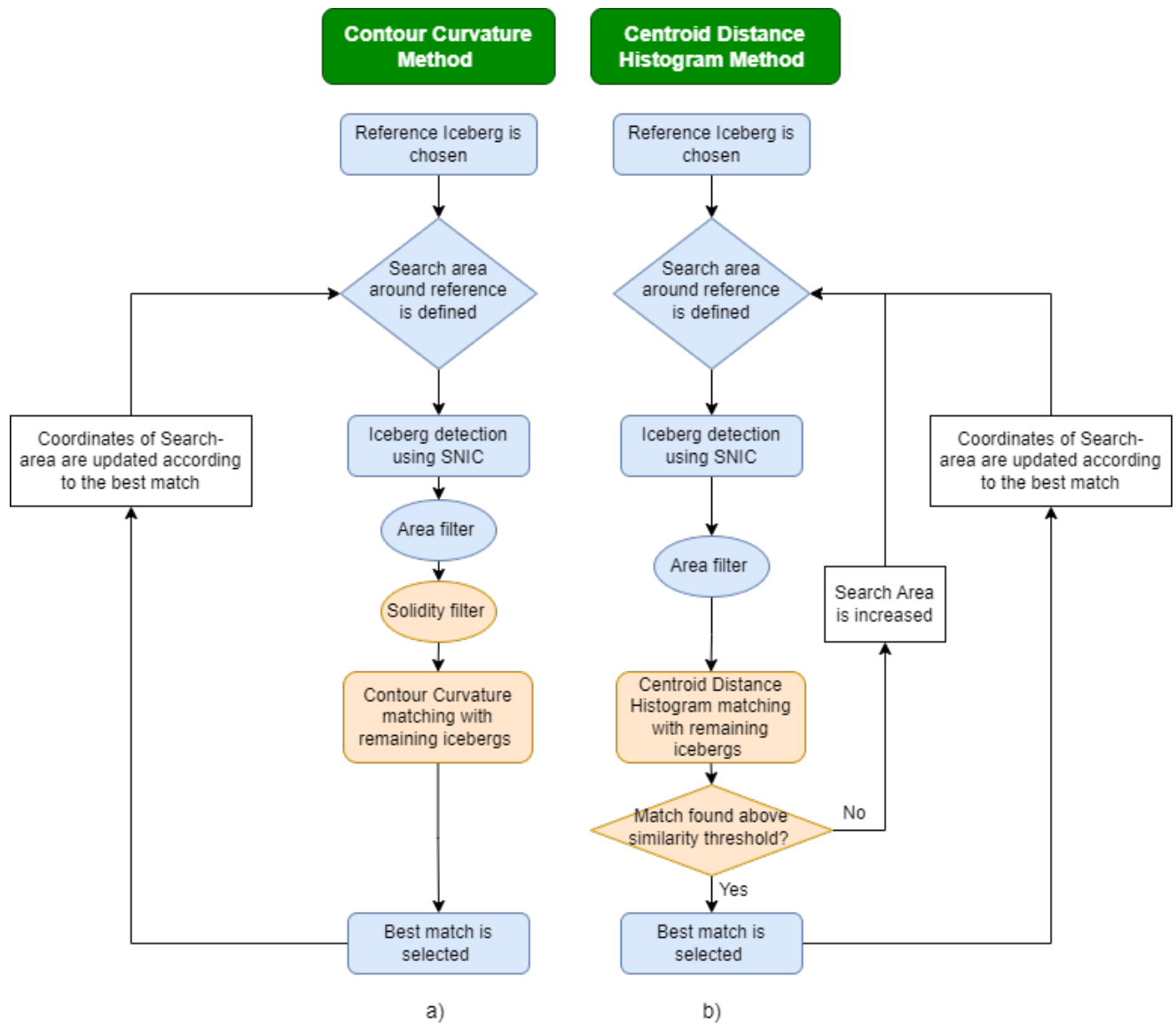
Figure 3.1: a) Schematic overview of a) the Contour Curvature method and b) the Centroid Distance Histogram method. In blue the steps which are equal, in orange the elements which differ for both methods.

3.1. Contour Curvature Method

First, a reference iceberg must be chosen for tracking. The goal of the algorithm is to recognize this iceberg automatically in subsequent time frames. This is accomplished by comparing the contours of the reference and potential target icebergs. Before this contour is extracted, icebergs need to be distinguished from their surroundings in the SAR images using a Simple Non-Iterative Clustering algorithm (SNIC) (Koo et al., 2021). The potential targets are first filtered on area and solidity, a shape property of the iceberg, to decrease the number of potential targets. Following, the curvature function of the reference and target icebergs are defined and the contours are divided into multiple segments. The individual segments of the reference iceberg are compared to those of target icebergs and the best match is selected. This process is summarized in Figure 3.1a. Each step is in greater detail described in the sections below.

Each of the steps below is performed for icebergs within the search area, which is defined as as a circle with a radius of 125 km centered on the coordinates of the reference iceberg. The steps are repeated once a day, for each day that SAR imagery is available. When a matching target is selected, the coordinates of the search area are updated automatically to match the centroid of this target.

3.1.1. Iceberg Detection Using Superpixel Segmentation

A superpixel segmentation method known as Simple Non-Iterative Clustering (SNIC) is used to identify icebergs in SAR images. This is a clustering method in which superpixels are grown from a grid of seeds, forcing connectivity from the beginning (Koo et al., 2021). Google Earth Engine includes SNIC as a built-in function. The number of generated superpixels, and thus the resolution at which icebergs can be resolved, is determined by the seed distance. A threshold function is used where superpixels with an average backscatter intensity above the threshold are classified as icebergs and superpixels below the threshold are classified as non-icebergs. Subsequently, adjacent iceberg superpixels are combined to form a single iceberg polygon (Koo et al., 2021). Before generating the superpixels, the data is smoothed with a Gaussian filter to reduce possible errors due to speckle effects in the SAR images (Koo et al., 2021).
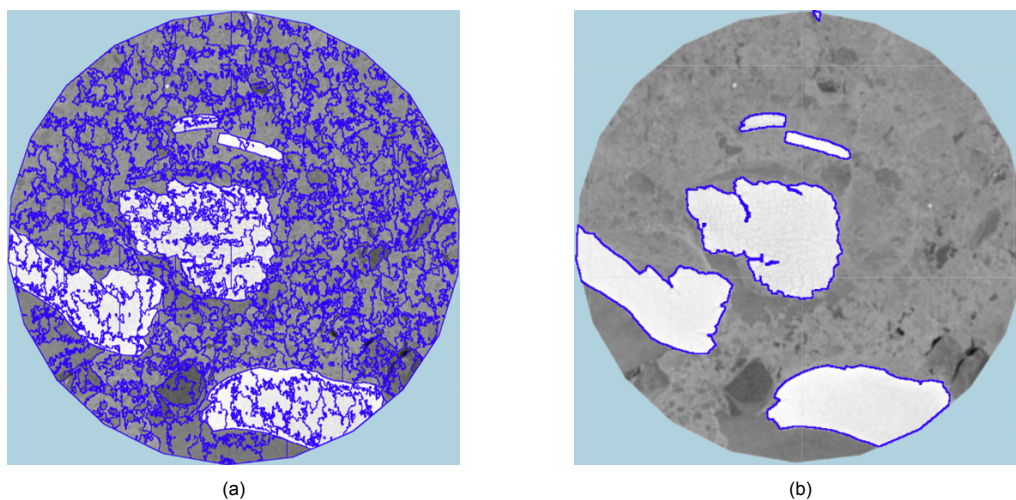

Figure 3.2: a) Segmentation into superpixels performed by SNIC (seedsize = 120). b) The superpixels with a brightness above the threshold are joined into iceberg vectors.

3.1.2. Filtering

Two filters are used before the matching step to reduce the total number of target icebergs in the search area before matching. The two filters are the area and the solidity filter. The goal of the area filter is to reduce the run time by eliminating unlikely matches that are much smaller or much larger than the reference iceberg. Simultaneously, a wide range of areas is allowed to ensure that the correct match is not eliminated if the surface area is reduced due to the iceberg being only partially visible, or enlarged due to adjacent smaller icebergs. The minimum of the area range is set to be the reference area divided by 2.5 and the maximum is set to the reference area multiplied by 2.5. The solidity filter is a shape filter

that reduces the chances of incorrectly matching a target. It also eliminates irregular shapes originating from snow covered land or bright sea ice areas, which frequently produce contours with a solidity much lower than the tracked icebergs. The allowed solidity range is $[S_{ref} - 0.05; S_{ref} + 0.05]$, where S_{ref} is the solidity value of the reference iceberg. The solidity S of a contour is a measure for cavities along the boundary and is defined as the contour area divided by the contour's convex hull area (see Equation 3.1). The convex hull can be described as the contour's perimeter as if it were surrounded by a rubber band (Olson, 2011). Figure 3.3 shows two iceberg contours (in blue) and their convex hull (in red). The iceberg shown in (a) has a much lower solidity compared to (b) due to its many cavities. In general, smooth, round shaped icebergs, which roughly approximate a circle or ellipse, have a solidity approaching one. Whereas irregularly shaped icebergs, with many cavities have a much lower solidity (Olson, 2011).

$$S = \frac{Contour\,Area}{Convex\,Hull\,Area} \tag{3.1}$$

Figure 3.4 displays an example of how targets are eliminated by filtering. The contour highlighted with red dashes, is selected as reference iceberg. All detected icebergs within the search area are highlighted in Figure 3.4a, it can be seen that also many land patches covered in snow and ice along the right edge of the search area are labelled as icebergs. Figure 3.4b depicts the remaining icebergs after filtering on area and Figure 3.4c depicts the two remaining targets after also filtering on solidity.
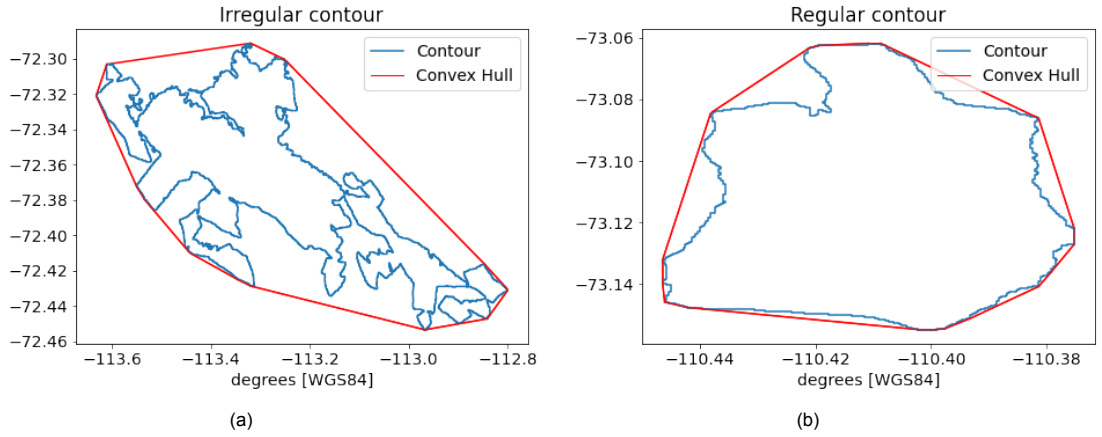

Figure 3.3: Contour (blue) and Convex Hull (red) of (a) an irregularly shaped contour with a solidity of $S = 0.518$ and (b) of a more regular shaped contour with a solidity of $S = 0.872$.

3.1.3. Curvature Function and Contour Segmentation

After filtering, this step is used to select the best match from the remaining targets. The curvature function is used to divide the iceberg contours into segments and compare the shape of these segments. The principle of comparing contours based on their curvature is derived from the methods described by Liu and Srinath, 1990, where the curvature function and segmentation are used for the partial recognition of shapes, and Da Gama Leitão and Stolfi, 2002, where the curvature function is exploited for ancient pottery fragment matching. The following steps are used to carry out segmentation and matching. First, the contours (of the reference and targets) are smoothed using a Gaussian filter, since the extracted contour is considered to be a noisy signal of the original iceberg due to inconsistencies in acquisition, conversion from the iceberg into pixels and changes over time due to wear and melt. Second, the curvature function of the contours is determined. The curvature function is a shape measure, that returns a high value for sharp bends and corners in the contour, while a flat line returns a curvature value of zero. The curvature function $k(t)$ is defined as follows (Liu & Srinath, 1990);

$$k(t) = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{[\dot{x}^2 + \dot{y}^2]^{3/2}} \tag{3.2}$$
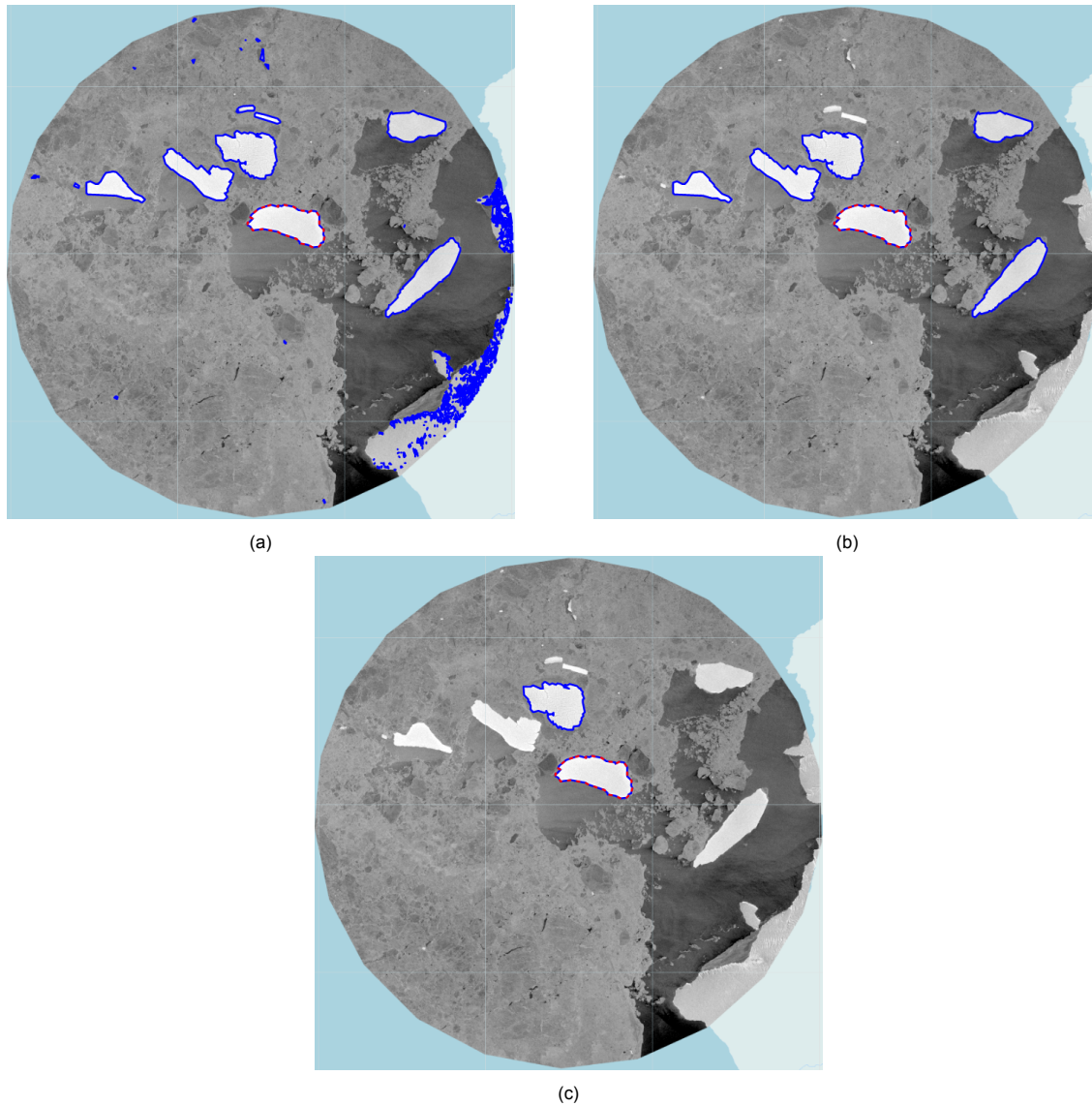
Figure 3.4: The effect of different filters on the amount of target icebergs in the search area. (a) All icebergs detected by SNIC within the search area are highlighted. (b) Icebergs remaining after filtering on area. (c) Icebergs remaining after filtering on both area and solidity. The iceberg which was chosen as reference iceberg is highlighted in red.

The discretization of the gradients in the curvature function was done using the *NumPy gradient* function in python (NumPy, n.d.).

Third, the contour and corresponding curvature function are divided into different segments at the curvature's minima and maxima positions. Figure 3.5a depicts a curvature function with the minima and maxima highlighted in red. Figure 3.5b shows the corresponding (smoothed) contour that has been segmented in the highlighted locations of the curvature function. The number of segmentation positions is dependent on the minimum absolute peak height, the prominence value (width of peak base) and the smoothing factor. The values chosen for the minimum peak height, prominence value and smoothing factor are 125, 100 and 1.0 respectively. These values were determined through a sensitivity analysis which is described in Appendix A. In the next step, the segments of the targets are compared to the segments of the reference iceberg to determine the best match.

3.1.4. Matching of Segments
The following steps are carried out for each target in order to determine the best match. 1) The error between the curvature function of each target segment and each reference segment is determined
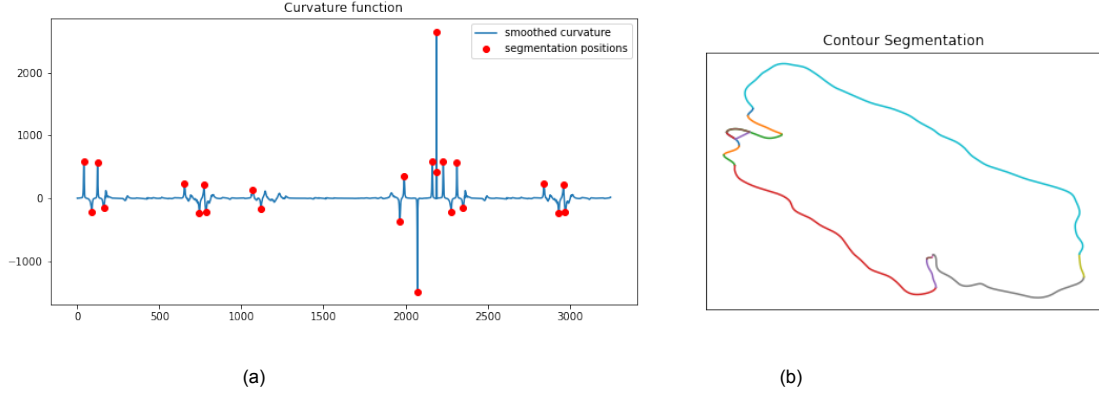
Figure 3.5: (a) The curvature function of the contour, the peak locations at which the contour is split are shown in red. (b) The smoothed contour, divided into different segments based on the curvature function.

using a method known as Dynamic Time Warp (DTW). DTW is a technique for comparing two series that allows for local stretching and compression of the series, as well as relative shifting of the end-points (Giorgino, 2009). Natural variability in iceberg contours is caused by wear, melt and image distortions, resulting in local distortions or elongations of the curvature function. By locally stretching and compressing the curvature function, DTW can compensate for these distortions and elongations. Figure 3.6 illustrates how stretching and compression can be achieved by connecting points of the blue curve to multiple or displaced points on the black curve. After stretching/compressing the two signals, the error is computed by summing over the distances between the curves along the connected points (Giorgino, 2009). 2) A reference-target segment pair is labelled as valid match if the error determined in step 1) is below a certain threshold. 3) The total number of valid segment matches in the target is counted. 4) The target with the most valid matches is chosen as best match. 5) If there are two targets with an equal number of valid matches, the contour with the lowest average error is chosen as best match.

6) The location of the search area is updated to match the centroid of the best match, and the process is repeated for SAR imagery of the next day. Furthermore, the reference contour needs to be updated occasionally, if the overall shape of the iceberg has altered too much due to disintegration or wear.

The method becomes rotation and translation invariant by employing the curvature function and segmentation. In addition, segmentation induces partial contour recognition, allowing for the tracking of disintegrated or partially visible icebergs.
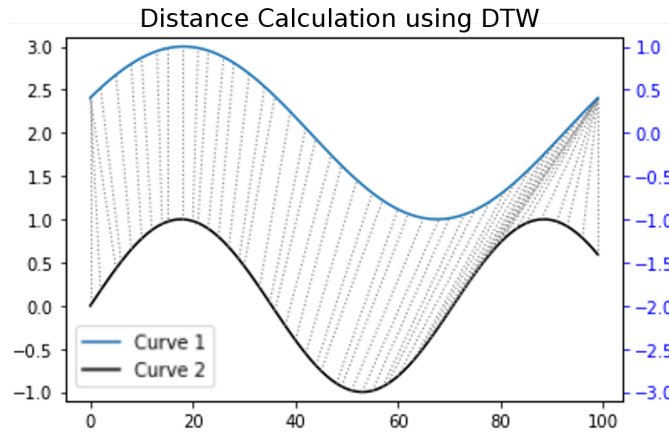

Figure 3.6: Matching of two functions using Dynamic Time Warping, the functions shown in blue and black, the dotted lines in between represent how each element from one function is linked to the other function. (This image was generated using the *FastDTW* python package based on the study by Salvador and Chan, 2004.)

3.2. Centroid Distance Histogram Method

The Centroid Distance Histogram (CDH) method developed by Koo et al., 2021 is described in this section. The initial steps, including Simple Non-Iterative Clustering (SNIC) and iceberg classification, are the same as described in subsection 3.1.1. The methods diverge after the potential target contours are extracted from the SAR imagery and the subsequent steps for the CDH method are described below. A few minor changes have been made to the algorithm described by Koo et al., 2021 to make it suitable for comparison with the CC method.

3.2.1. Filtering

The icebergs are filtered based on their area, in a similar way as described in subsection 3.1.2 for the CC method. The area filter in this method, however, employs a smaller range and allows targets with areas between 0.8 and 1.2 times the reference area. The effect of this smaller area range is that the throughput after area-filtering is much lower compared to the CC method.

3.2.2. Centroid Distance Histogram

A centroid distance histogram is generated for the reference iceberg and each of the remaining targets after filtering. This histogram serves as a unique function that represents the iceberg's shape and area and is used to determine the similarity between the reference and targets. To find the best matching target, the following steps are executed. 1) A centroid distance histogram is generated for the reference and the different targets. This histogram is made by measuring the distance between all pixels within the iceberg and the iceberg's centroid. A histogram is then constructed of the occurrences of all distances. This histogram therefore represents the overall shape and area of the iceberg (Koo et al., 2021). 2) The histogram of the target is compared to the reference histogram using the similarity function. The similarity S between the reference R and target T is defined as follows;

$$S(T,R) = \left(1 - \frac{\sum_{i=1}^{m} |N_T(i) - N_R(i)|}{\sum_{i=1}^{m} N_R(i)}\right) \tag{3.3}$$

In this formula m is the bin size of the histogram and $N_R(i)$ and $N_T(i)$ are the number of pixels in the ith bin for the reference and target iceberg respectively (Koo et al., 2021). Each bin represents a pixel-centroid distance range.

4) Targets which have a similarity score greater than 0.7 are considered to be a match; if there are multiple targets with a similarity score greater than this threshold, the target with the highest similarity score is chosen as best match. 5) As in the CC method, the location of the search area is updated to match the centroid of the best match and the process is repeated for SAR imagery of the next day. If there is no SAR imagery available, the search area is extended with 2.5 km for each day without data. As in the CC method, the reference contour needs to be updated occasionally, if the overall shape of the iceberg has altered too much due to disintegration or wear.

To allow for a valid comparison of the CDH and the CC method, the CDH algorithm has been slightly modified in relation to the steps described by Koo et al., 2021. Table 3.1 provides an overview of these adaptations, which include the following. The search area is expanded so its size matches the search area used in the CC method, this extension was also required to make the method suitable for the tracking of large icebergs, which would otherwise not fit in the initial search area. The increment of the search area is reduced as the overall search area is already enlarged. Furthermore the similarity threshold was reduced to 0.7, to allow for less frequent updating of the reference contour.

3.3. Comparison Between CC and CDH Method

For benchmarking, the performance of the CC method is compared to that of the CDH method. Both methods are used to track the same 15 icebergs over the course of one year. The reference contour remains constant throughout the tracking period and is only updated if the overall shape of the iceberg has changed significantly due to natural decay. For each of the tracked icebergs, the number of required manual interruptions, as well as the recall is registered. The recall (Equation 3.4) is calculated by dividing the number of correctly selected matches, by the number of satellite images in which the

CDH method (Koo et al., 2021)	Adapted CDH method
Reference polygon is manually drawn	Reference polygon is imported as contour found with the SNIC method
Search area with 25 km radius	Search area with 125 km radius
Search area is increased with 25 km if no match is found	Search area is increased with 2.5 km if no match is found
Minimal similarity threshold of 0.8	Minimal similarity threshold of 0.7

Table 3.1: Adaptations made to the CDH method described by Koo et al., 2021 in order to make it more suitable for a variety of iceberg sizes and for comparison with the CC method.

iceberg was for at least 40% visible. Visual inspection is used to determine whether the method returns the correct match for each time step, as deciding whether a match is false or correct is a simple task for humans.

$$Recall = \frac{Nr.CorrectMatches}{Nr.AvailableImages} \tag{3.4}$$

4

Results

The 15 icebergs described in section 2.4 were tracked for one year using both the Contour Curvature (CC) method and the Centroid Distance Histogram (CDH) method as described in section 3.1 and section 3.2, respectively. An overview of the trajectories found by both algorithms is shown in Figure 4.1. The beginning of each iceberg trajectory is displayed in blue and transitions to red throughout one year. The locations of manual interruptions are indicated with circles for both the CC method (in green) and the CDH method (in pink). The results obtained with both algorithms are presented in section 4.1 and section 4.2, after which the performance of both algorithms is compared in section 4.3. In this chapter an overview of the most important results is presented. For an overview of the exact tracking results of the 15 individual icebergs, Appendix B can be consulted.
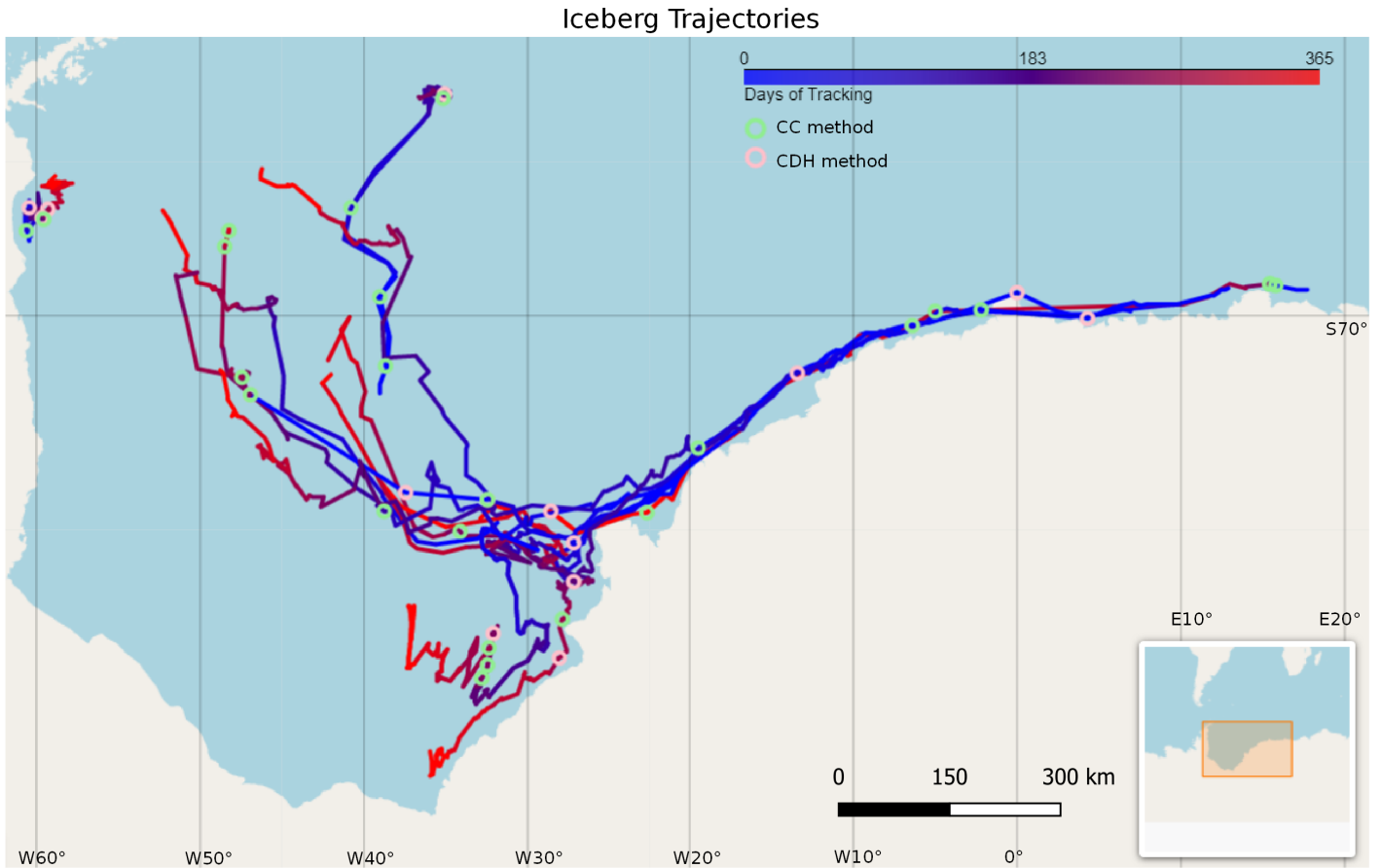
Figure 4.1: Trajectories of the tracked icebergs. Each iceberg trajectory starts as blue and progresses to red during one year of tracking. Manual interruptions needed while tracking are shown in pink for the CDH method and in green for the CC method.

4.1. Contour Curvature Method

Figure 4.1 shows the found iceberg trajectories. A clear difference in the divergence of the trajectories can be seen at 20°W. East of 20°W the trajectories follow essentially the same path, which is due to the strong Antarctic Coastal Current in this area. West of 20°W, the icebergs reach the Weddell Gyre and the trajectories diverge more from one another. The greatest distance travelled by an iceberg is about 6930 km. Occasionally, a manual interruption of the method is required, for example, when the iceberg drifts outside the search area. The places where these manual interruptions were done are marked with green circles in the trajectories in Figure 4.1.

An example of the trajectory of an iceberg is shown in Figure 4.2. Again, the beginning of the trajectory is shown in blue and the end of the trajectory is shown in red. In total, the iceberg was correctly selected 82 times, while it was visible in 133 SAR images over a period of one year. Therefore, the recall is 0.62 for this iceberg. There are several reasons why the iceberg is not always correctly selected, which are discussed in more detail later in this section. This iceberg roughly follows the pattern of the Weddell Gyre.

In Figure 4.3, six stills of the tracking process are shown for the tracking of a very large iceberg with an area of about 5600 km^2. The SAR data available for the given day is displayed in the search area with a radius of 125 km. The location of the search area depends on the center of the previously selected target. In green the selected target of the corresponding day is shown, in light green all previously (correctly) selected targets are visible. Figure 4.3b, 4.3c and 4.3e show examples of situations where the iceberg is partially visible within the SAR image. Figure 4.3b and 4.3d illustrate how the presence of meltwater (4.3b) or sea ice (4.3d) distorts the detected iceberg contour.
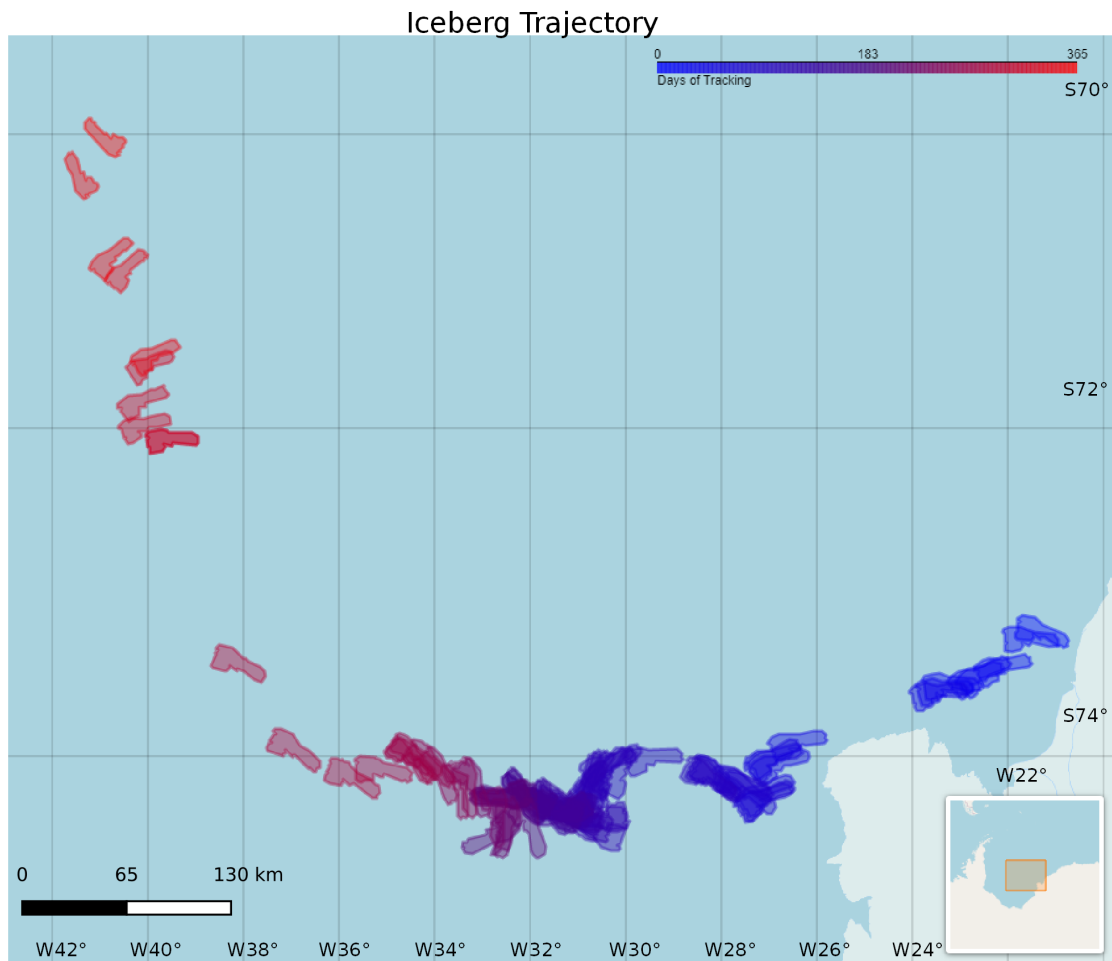
Figure 4.2: Example trajectory of an iceberg with an area of 450 km² over a period of one year, blue indicating the start of the tracking period, red indicating the end of the tracking period.

Several relations between the tracking performance and iceberg properties are apparent when tracking the 15 icebergs. The performance, expressed as recall, depends to some extent on the area of the respective iceberg, as illustrated in Figure 4.4a, where the results for the CC method are plotted in blue circles. For one very small iceberg (3 km²) the performance is poor, as represented by the leftmost data point in Figure 4.4a. For icebergs with areas between 16 and 560 km², the recall ranges from 0.24 to 0.88. The best results are obtained for areas between 66 and 239 km², although it is not possible to say whether this higher performance is solely due to the size, since other factors are also involved. For one very large (5600 km²) iceberg, the performance is around 0.4.

(a) Day 0   19-11-2017

(b) Day 38   27-12-2017

(c) Day 80   07-02-2018

(d) Day 145   13-04-2018

(e) Day 293   08-09-2018

(f) Day 348   02-11-2018
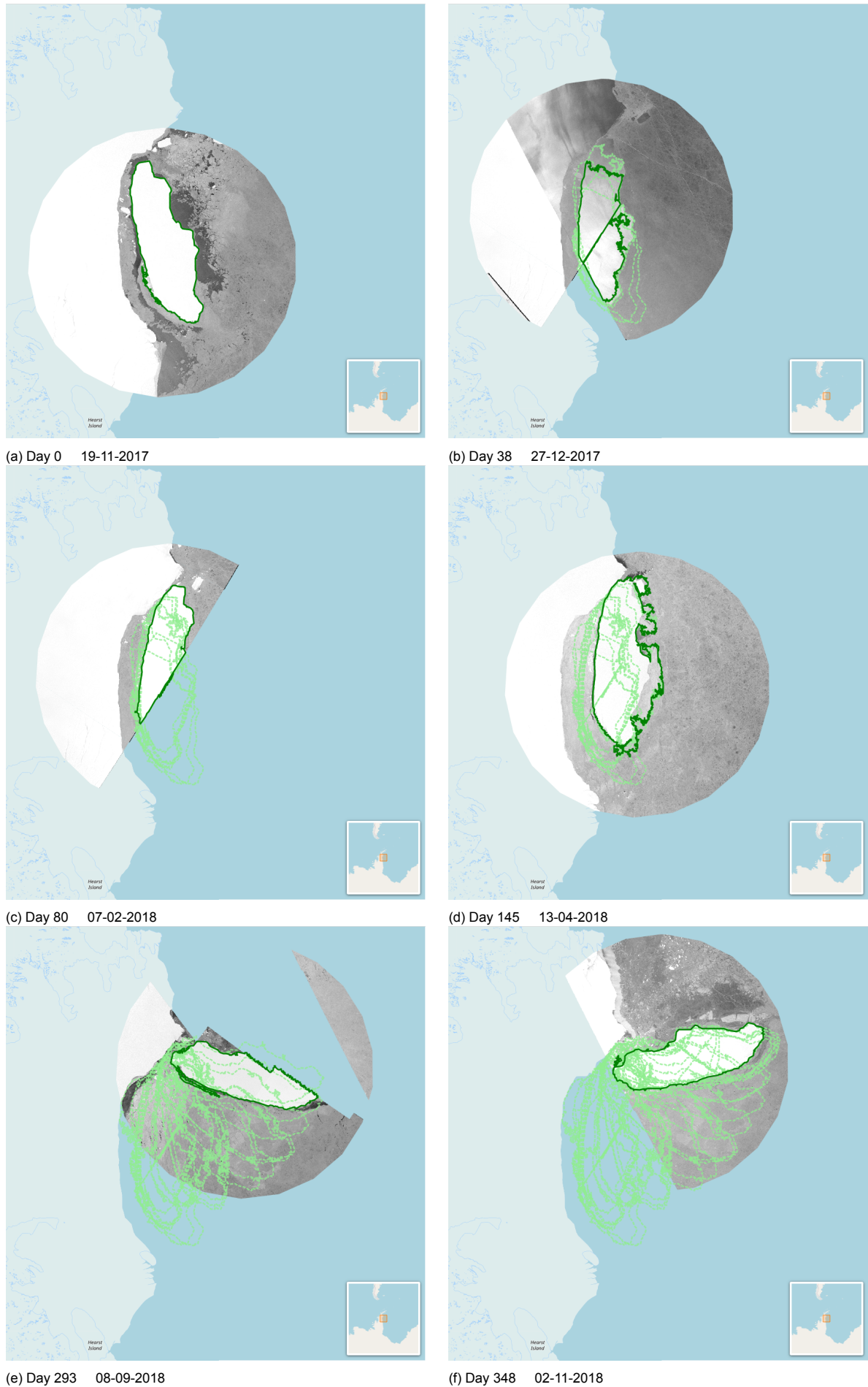
Figure 4.3: Iceberg with an area of approximately 5600 km^2, selection of 6 contours found over different days within the tracking period of one year. For each day the SAR data within the search area is shown, the location of the search area is dependent on the centroid of the previously selected target. In green the selected contour found on that day is shown, in light green the contours of previously selected targets, wrong matches omitted.

(a)

(b)

(c)
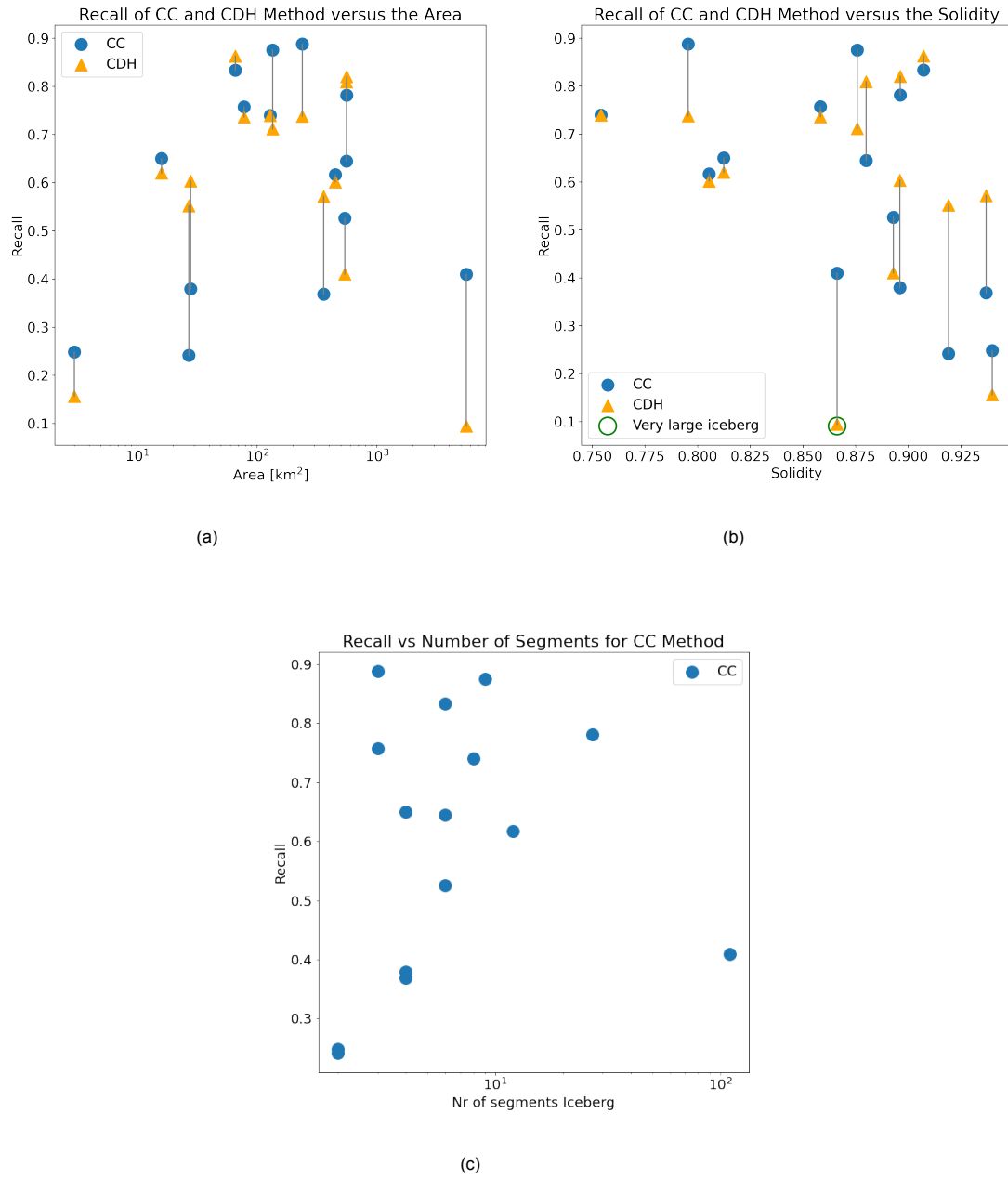
Figure 4.4: Coupled scatter plot of the CC and CDH method showing the recall as a function of (a) the area [km^2] of the iceberg and (b) solidity of the iceberg. Each blue or orange point represents a tracked iceberg, the gray lines join the outcomes of the same iceberg for the two different methods. (c) Recall of the CC method versus the number of segments within the contour.

Figure 4.4c shows that for most icebergs, a positive correlation can be found between the number of segments in the contour and the performance. The reason for this is that the contours with many segments are likely to have a more recognisable shape. This effect can also be seen in the blue circles in Figure 4.4b, where a general trend of decreasing performance is observed for icebergs with higher solidity values. A higher solidity value means that these icebergs often have simpler shapes that are more elliptical or rectangular. The relation between solidity and the number of segments is shown in Figure 4.5, where it can be seen that there is a negative relation between the two values. In general, icebergs with a higher number of segments have lower solidity and these icebergs are easier to track than icebergs with higher solidity and a lower number of segments.
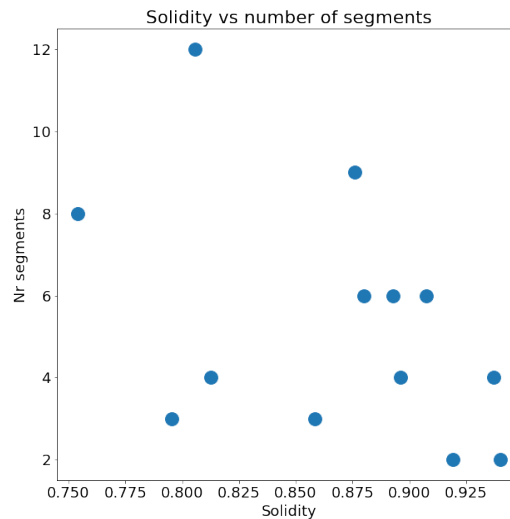
Figure 4.5: The solidity of the icebergs versus the number of segments in the contour. For these icebergs, a negative relation can be seen. Two large icebergs were omitted for clarity.

More than half of the tracked icebergs, required one or more interruptions of the algorithm within the tracking period, caused by flaws in the algorithm or in the available data. The locations of manual interruptions for the CC method are indicated by green circles in Figure 4.1. Two situations occur frequently where manual interruptions are required. The first is when the iceberg has moved out of the search area. This situation occurs when the iceberg drifts too far before new satellite images are available in the following days, either due to a large drift velocity of the iceberg or due to a large data gap in a given time period. Large drift velocities are observed mainly along the Antarctic Coastal Current; in Figure 4.1 the icebergs to the right of 20°W closely follow this current. Another situation where the iceberg appears outside the search area, is when a target with a large distance to the correct match was selected in the previous time frame. The search area is then updated with the coordinates of the incorrect target, causing the search area to shift. The second reason for manual interruptions is when the iceberg contour of the iceberg being tracked is altered too much due to melt and wear and an update of the reference contour is required. The average number of interruptions is 1.9 for this method.

In general, a large proportion of correctly found targets result from the combination of the area and solidity filter. Of all the correct targets found for the 15 icebergs, about 26% were found by the algorithm's matching step, while in the other 74% the target was automatically selected, passing as the only target after filtering for area and solidity. Moreover, there is a negative trend between the performance and the percentage of how often the matching step was used to find the correct target. Of the situations where incorrect targets were selected, about 38% were incorrectly selected by the matching step, and the other 62% were selected while the correct target was not visible in the SAR image. Overall, incorrect targets were selected 23% of the cases. Thus, using the area and solidity filter provides an advantage for quickly selecting the correct match. However, using the solidity filter can also lead to exclusion of the correct target if its contour is altered by the presence of adjacent sea ice or smaller icebergs. An

example of this can be seen in Figure 4.6, where the solidity of the iceberg in (b) is significantly reduced by the presence of smaller iceberg pieces that are included in the contour.

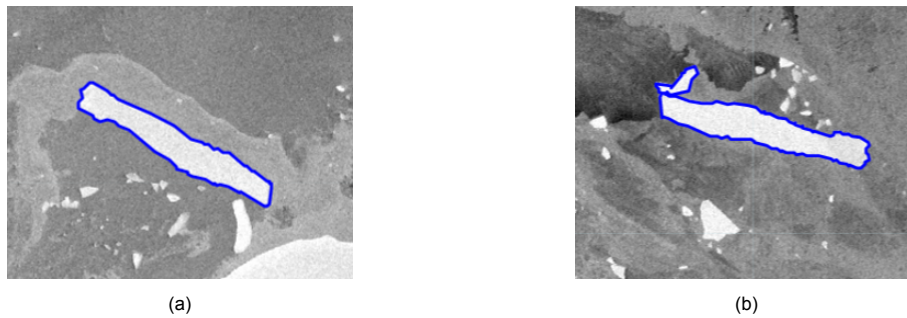(a)   (b)

Figure 4.6: An iceberg depicted at different moments in time. The solidity value is altered due to the presence of small icebergs, which prohibits the algorithm from identifying the correct match as the new solidity is now outside of the filter range of +-0.05. (a) Solidity = 0.8610 (b) Solidity = 0.6522

The area filter eliminates much larger and smaller icebergs compared to the reference, but still allows targets that are 2.5 times smaller or larger. Because of this large range, the filter does not often result in exclusion of the correct target. Exclusion due to the area filter only occurs when less than 40% of the iceberg is visible or when the iceberg is too close to an iceberg that is more than 1.5 times the size of the reference iceberg. The main advantage of the area filter is the significant reduction in computational steps, since the number of targets is significantly reduced. A second advantage is the reduction of the probability of an incorrect match due to the smaller number of targets.

4.2. Centroid Distance Histogram Method

The same 15 icebergs described in section 2.4 are tracked using the Centroid Distance Histogram (CDH) method developed by Koo et al., 2021. Again, the relations between iceberg characteristics and the tracking performance are examined. Figure 4.4a displays that the performance of the CDH method is poor for very small (<10 km^2) or very large icebergs (>1000 km^2). For small to large icebergs (>10 and <1000 km^2), performance is generally better. The best results can be obtained for icebergs between 50 and 250 km^2. However, the number of tracked icebergs is too small to conclude that this relation holds for all icebergs in this range, since there are other characteristics and circumstances that affect the performance.

The orange triangles representing the CDH method in Figure 4.4b show a weak relation between the iceberg solidity and the tracking performance. For icebergs with a lower solidity (<0.85), the performance was generally higher than 0.5. One iceberg (encircled in green) does not fit in this pattern. This point corresponds to the very large iceberg; why the performance for this iceberg is low is discussed later in this chapter. If this iceberg is omitted, a negative correlation between the solidity and performance can be seen.

Interruption of the tracking was occasionally necessary. The locations where this was necessary are marked by a pink circle in Figure 4.1. Interruption was especially necessary in situations where the iceberg had a high drift velocity. Another reason for interruption is the need to update the reference iceberg when its shape or area has changed too much due to melting or disintegration. On average, 1.0 interruptions were required per iceberg trajectory. Overall, wrong matches were selected in about 10% of the situations.

The area filter in the CDH method is relatively narrow, ranging between 0.8 and 1.2 times the area of the reference iceberg. This results in the exclusion of the correct target iceberg if more than 20% of the iceberg is outside the image, or if the iceberg is adjacent to sea ice or smaller icebergs that have an area greater than 20% of the reference iceberg. An advantage of this filter is that it effectively reduces the large number of targets that are similar in size, which speeds up the algorithm.

4.3. Comparison Between the CC and CDH Method

For small icebergs (<10 km^2) both methods show poor performance. When examining the recall for small to large icebergs (>10 and <1000 km^2), there is not a large difference between the Contour Curvature (CC) and the Centroid Distance Histogram (CDH) method. For four of the icebergs, the CC method shows higher performance, while for four other icebergs, the CDH method shows higher performance. For the remaining 7 icebergs, the recall is within 10% difference from the other method, so the performance is considered to be equal. Wrong targets are more often selected in the CC method compared to the CDH method. The CC method wrongly selects a target around 23% of the time while the CDH method has an average of 10% wrongly selected targets. A side effect of wrong-selected targets is the deviation in the location of the search area. If this deviation is large, the correct target can be located outside of the search area in the successive SAR images. This, consequently, results in the algorithm having to be interrupted to bring the search area back into the desired location.

| (a) | (b) |
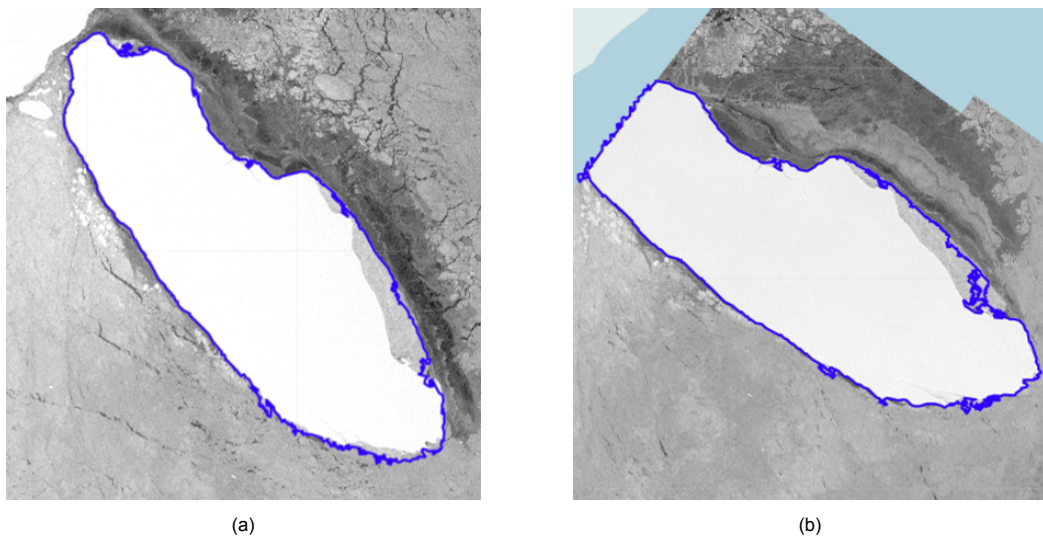
Figure 4.7: (a) Contour used as reference, obtained at 2018-08-01 where the iceberg is fully visible within the satellite image. (b) Contour used as target, obtained at 2018-08-09, the iceberg is only partially visible in the satellite image.

For the largest iceberg, which has an area of >5600 km^2, the difference in performance between the CC and CDH methods is large. For the CDH method the recall is 0.10 and for the CC method this is 0.41. The main reason for this difference is that the iceberg is often only partially visible in the SAR images due to its large size. For the CDH method, targets that are less than 0.8 or more than 1.2 times the reference area are not considered as target. If they are within this range, but a smaller part of the iceberg is not visible, this directly affects the placement of the centroid and hence the centroid distance histogram, as shown in Figure 4.8. Although the overall shape of the histogram is preserved, the similarity is reduced due to the shift of the centroid. If a large part of the contour is still visible, the CC method is able to match these segments with the corresponding reference segments, as shown in Figure 4.9, so that the correct target can still be selected. The reason that the recall for the CC method is still relatively low is mainly due to the presence of sea ice in the period from February to May 2018, which often makes it hard to distinguish the iceberg from its surroundings as the sea ice leads to large Figure 4.10 or smaller Figure 4.3d distortions of the contour.
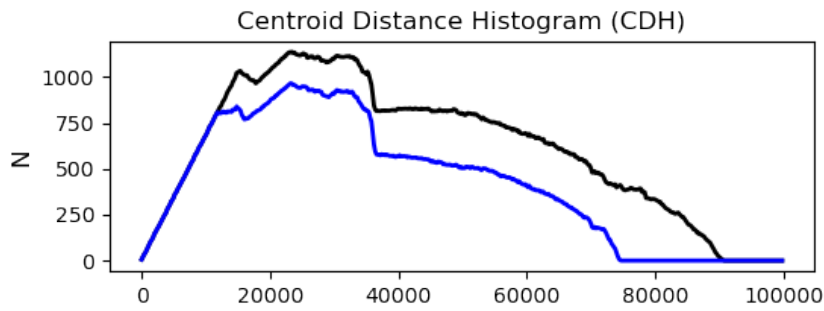
Figure 4.8: Centroid Distance Histogram for Iceberg 6 reference (black) and target (blue) as shown in Figure 4.7. The histograms have a similarity value of 0.68. The overall shape of the histograms remains the same, however, due to displacement of the centroid, the similarity is decreased.
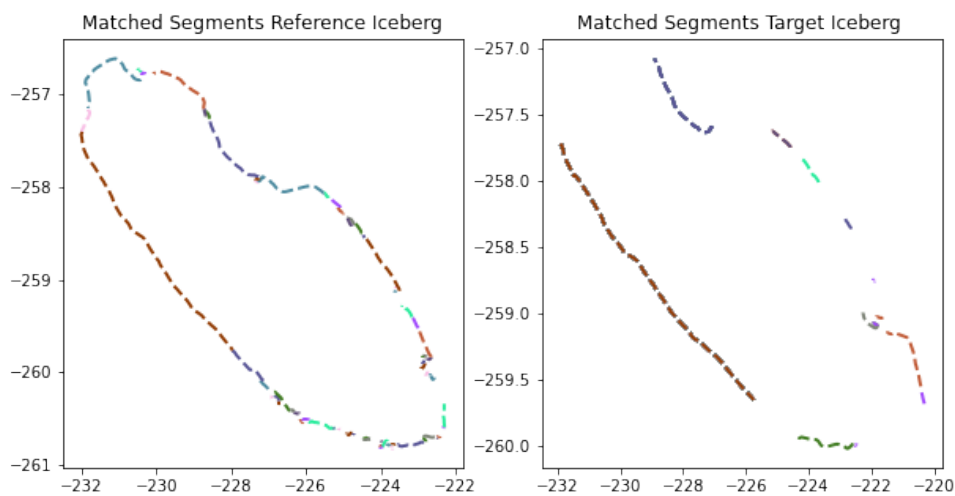

Figure 4.9: Matched segments by the CC method for iceberg 6 as shown in Figure 4.7. The reference and target contour contain respectively 184 and 256 segments of which 64 were matched to each other.

4.3.1. Implications for Both Methods

One disadvantage of both methods is that the size of the search area is limited by the computation limit of Google Earth Engine. This limit lies around a search area with a radius of 125 km, above this radius, the amount of computations that need to be processed is so large that the risk of exceeding the Google Earth Engine capacity becomes a limiting factor. This is a problem because sometimes there is a gap in the available data and the iceberg could be drifting outside the search area due to ocean currents. Although the CDH method is programmed to slightly increase the search area for each day that no iceberg match is found, sometimes this increase is not sufficient. If the iceberg is outside the search area, a manual interruption is required to update the location of the search area.

Another problem which arises for both methods lies within the superpixel segmentation step of the algorithm, which is the same for both the CC and CDH method. As the iceberg classification is based on a brightness threshold, this is in some situations malfunctioning when the surrounding sea ice appears to bright in the SAR images. This situation is illustrated in Figure 4.10, where an iceberg is visible as a white structure in the middle surrounded by light grey sea ice. However, as the surrounding sea ice also exceeds the brightness threshold, adjacent areas are added to the contour of the iceberg, strongly distorting the contour shape and therefore also causing strong deviations in the area and solidity value. Depending on the size and shape of the added sea ice area, this effect can lead to exclusion of the iceberg as target in both the CC and the CDH method.
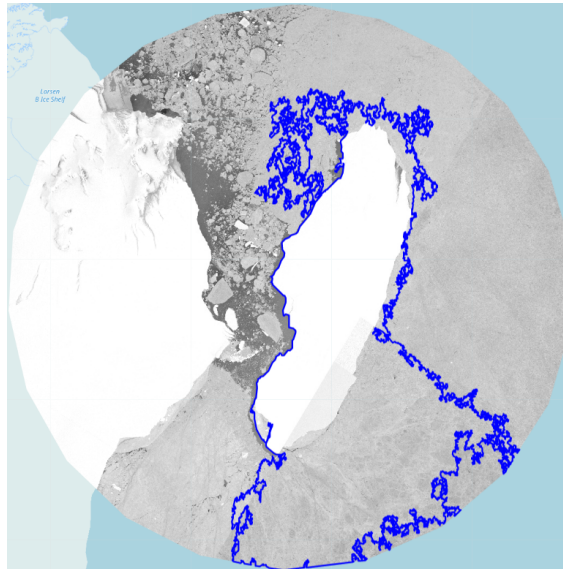
Figure 4.10: Surrounding sea ice gets identified as part of the iceberg because its brightness is above the threshold value. This effect strongly effects the contour of the iceberg and thus the area and solidity, therefore it becomes impossible to indicate the iceberg as correct match.

5

Discussion

In this study, a new tracking method is presented based on contour recognition, including filtering of targets based on area and solidity and subsequently contour segmentation and matching based on the contour's curvature. The preliminary step of iceberg detection in Synthetic Aperture Radar (SAR) images is done using a pre-existing method based on Simple Non-Iterative Clustering (SNIC) (Koo et al., 2021). This new method, called the Contour Curvature (CC) method, enables partial contour recognition, allowing to track icebergs which are only partially visible within SAR images. The CC method is compared to the already existing Centroid Distance Histogram (CDH) method, developed by Koo et al., 2021 to benchmark the performance.

5.1. Performance of the Method

The performance of the method is dependent on several properties of the icebergs and their surroundings. First, the performance is dependent on the solidity and number of segments in the iceberg contour, both are indicators of the complexity of the shape. In general, a lower solidity and higher number of segments imply more distinguishable iceberg shapes that are more effectively tracked. Second, the performance of the method is dependent on the surroundings of the icebergs. For example in the case of surrounding sea ice which appears bright in SAR images, the iceberg can not be (fully) distinguished from its surroundings anymore. Also, meltwater present on the surface of the iceberg can lead to a darker backscatter signal, prohibiting the iceberg to be detected.

Furthermore, the performance of the CC method is influenced by the used area and solidity filter. These filters exclude a large part of the iceberg targets, leaving only a few or exactly one target for the matching step, effectively reducing the possibilities for wrong matching. Frequently, solely one target remains after filtering, this target is then automatically selected as match. An advantage of this is that it speeds up the method as the matching step is omitted. However, this remaining target can also be a faulty match, which often happens in situations where the correct target is not visible in the SAR image, or the correct target is excluded by the solidity filter. The latter happens when adjacent sea ice or icebergs enlarge the contour of the iceberg, altering its solidity value.

To benchmark the performance of the CC method against an already existing method, a comparison with the Centroid Distance Histogram (CDH) method is made as described by Koo et al., 2021. Both methods perform poorly on very small icebergs, due to the large number of icebergs with similar areas and similar overall shapes, leading to more opportunities for wrong matching. For icebergs between 10 and 1000 km^2, the recall is similar, with one method occasionally outperforming the other. Very large icebergs with lengths in the range of the SAR swath width (250 km (ESA, n.d.)) are often only partially visible in SAR images. For these icebergs, the CC method clearly outperforms the CDH method. For the CC method, the combination of contour segmentation and the wide area filter effectively allows tracking of icebergs that are only partially visible. If part of the iceberg is not visible, this has a strong impact on the computed CDH and thus significantly reducing the performance of the CDH method.

5.2. Limitations and Suggestions for Future Research

The newly constructed CC method has some fundamental limitations. If the correct target is not visible, the CC method regularly selects the wrong target. One way to reduce the occurrence of this problem is to collect SAR images over several days until the search area is filled to a large extent (e.g. 80%), as this makes it more likely that the correct target is visible in the search area. The reason this approach was not implemented in this study is that a valid comparison between the CC and CDH method requires the first steps of both methods to be equal. To further reduce the number of falsely selected targets, two improvements are needed. First, the performance of the matching step needs to be enhanced, by further limiting the allowable stretch and compression by the Dynamic Time Warp error function.

Second, targets are often automatically falsely selected in the absence of other targets after filtering. To reduce this kind of wrong matches, a threshold function is needed before automatic selection. This threshold can be in the form of a similarity value as described for the CDH method or as a minimum number of matched segments before an iceberg is selected as a match. In addition, a widening of the solidity filter allows tracking in situations where the contour is distorted by adjacent smaller icebergs. However, this is only possible if the matching step has been improved to prevent the increase of faulty matches by widening the filter.

It is unlikely that this method can be fully automated in the future, as manual interruptions are required. Such as occasionally updating the reference contour when the shape has significantly altered due to melt and wear. Besides that, interruptions of the method are regularly needed when the iceberg drifts outside of the search area, while further enlargement of the search area poses problems concerning the GEE capacity and run time.

One proposal to resolve the latter problem is to introduce a multiscale method. In this study, the used resolution and SNIC grid size were kept constant for all tracked icebergs. Resampling the images to a lower resolution in combination with a larger seed spacing in the SNIC step, results in small icebergs not being detected, leading to fewer targets within the search area. The difference between detected icebergs by different seed spacing is demonstrated in Figure 5.1. A reduction in targets before filtering, reduces the computational time and complexity and allows for the enlargement of the search area.

Another technique to prevent the iceberg from getting out of sight is to include known directions of the ocean currents. Instead of enlarging the search area, the search area can be reshaped to match the likely drift direction of the iceberg based on the general ocean currents. This could be helpful in areas with strong drift which has a constant direction, such as along the Antarctic Coastal Current. However, In other areas where the current is less strong and does not have a constant direction, this might not work.

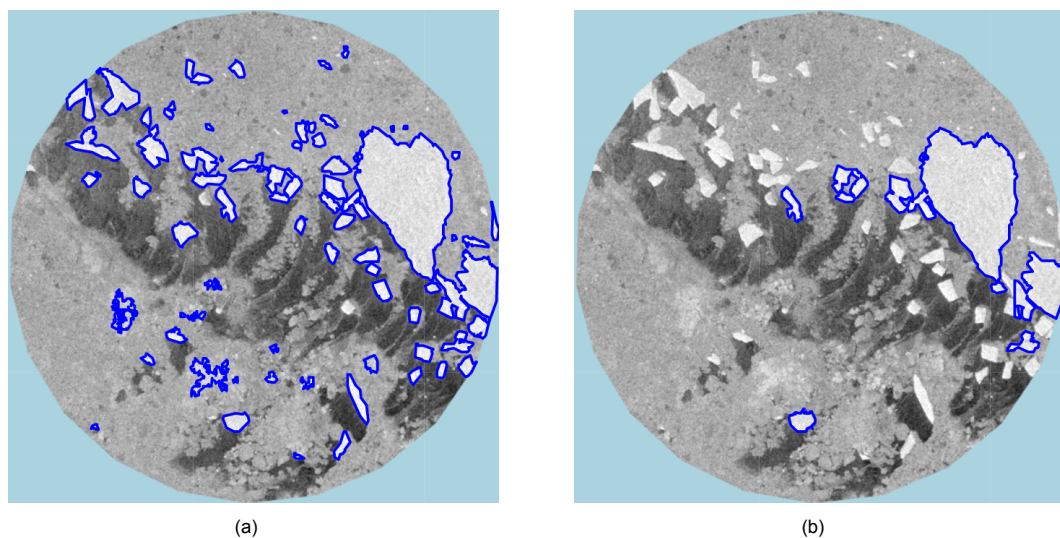(a)   (b)

Figure 5.1: Detected icebergs using SNIC with different seed distances in pixels (a) seed distance is 40, 6 icebergs are detected (b) seed distance is 200, 70 icebergs are detected.

5.2.1. Combining the CC and the CDH Method

For future studies, a suggestion is to combine the CC and CDH method. Aspects of the two methods can be combined in multiple ways. First, the CC method can be used with an additional similarity threshold after the segment-matching step. This can provide a solution for automatic selection of wrong targets when the correct target is not visible.

Second, the performance for small icebergs (<10 km^2), can be improved by using the CDH method in combination with an additional solidity filter. For small icebergs, it was found that the CDH method often selects a false target that is very close to the reference in area but with much lower solidity. While the CC method often selects a wrong target that is similar in solidity but has a larger difference in area. The use of a narrow area filter and the inclusion of a solidity filter, results in the exclusion of the iceberg targets which are often wrongly selected as a match. Since these small icebergs are in general fully visible within one SAR image, a narrow area filter does not lead to exclusion of the correct target. And since icebergs in this area range are abundant, stricter filtering can be particularly useful in this situation. For larger icebergs, this combination still poses problems if an iceberg is only partially visible.

A third possibility is to adapt the similarity calculation in the CDH method by using Dynamic Time Warping as used for segment matching in the CC method. In Figure 4.8 is demonstrated that for partially visible icebergs, the general shape of the histogram remains similar to the original histogram, but an overall shift in both the distance and number of occurrences can be seen. When using Dynamical Time Warp to calculate the error between the reference and target histogram, these shifts can be accounted for. However, this method needs to be investigated further, as allowing for stretching and shifting of the histogram can also favour the selection of wrong targets. In addition, the area filter should be widened so partially visible icebergs are not excluded.

5.2.2. Other Iceberg Detection Methods

One negative effect which occurs for both methods is the contour distortion of the detected icebergs due to the presence of sea ice or meltwater (Figure 4.10). The SNIC method in combination with a threshold does not allow for enough flexibility for these situations with varying backscatter. In recent years, other detection methods have been introduced that can detect icebergs in these situations with higher accuracy. One of these methods is described by Barbat, Wesche, et al., 2019, where a machine learning technique is used to detect icebergs around the Antarctic continent. Due to its adaptive nature, this method is successful in detecting icebergs under various backscatter conditions. In future studies, the SNIC method could be replaced by this method to allow for better performance. However, this method is computationally more complex and has not been run using Google Earth Engine.

The CC method shows that for partially visible icebergs, (partial) contour recognition is better for tracking compared to methods that depend on centroid placement (Barbat et al., 2021; Koo et al., 2021). In future studies, other contour recognition methods that allow for partial shape matching can be investigated for iceberg tracking. In addition, this study demonstrates how Google Earth Engine (GEE) enables the effective processing of large datasets. The use of GEE is recommended for future studies where large processing amounts of data poses problems for local computing.

5.2.3. Applications of the Tracking Results

The tracking results obtained from this study are suitable for various applications. The generated trajectories can be used to study ocean currents and general iceberg drift patterns, as well as for mapping of fresh water input locations. To estimate the contribution of fresh water volume from the tracked icebergs, the area reduction over time can be combined with altimeter measurements (Braakmann-Folgmann et al., 2021). However, because the contours are occasionally distorted by sea ice, adjacent icebergs or meltwater on top of the iceberg, some of the measurements are unreliable for calculating the fresh water volume. Furthermore, velocities and directions can be estimated from the known iceberg locations and their trajectories and used to inventory potential hazards to ship navigation and offshore projects.

6

Conclusion & Recommendations

A new method for automatically tracking icebergs has been developed and tested for icebergs in the Weddell Sea. The method, called Contour Curvature (CC) method, uses the curvature function of iceberg contours to find a chosen reference contour in subsequent SAR images. This contour and the corresponding curvature function, are divided into segments to enable the recognition of partially visible icebergs. The objective is based on four research questions, which are answered below.

What is the influence of different iceberg properties and surroundings on the performance of the method?
The method shows poor performance for small icebergs (<10 km^2) and divergent performance for medium to large icebergs (>10 and <1000 km^2). There is a negative relation between the solidity of an iceberg and the tracking performance and a positive relation between the number of segments in an iceberg contour and the tracking performance. a higher solidity corresponds to icebergs with simpler, more regular shapes. Icebergs with lower solidity and a greater number of segments correspond to more complex, better distinguishable shapes and thus better tracking performance. In addition, the SAR backscatter value of the iceberg and its surroundings have an impact on the detection and extraction on the iceberg contour. If the iceberg is surrounded by sea ice which appears very bright in the SAR image, this can lead to a distorted contour, since part of the sea ice is included in the iceberg. If there is melt water on top of the iceberg, it appears darker in SAR images and therefore may not be (fully) detected as an iceberg. Also, the presence of other icebergs with similar area and solidity to the reference, increases the chances of a false match.

How large is the contribution of the area and solidity filter on the ability to correctly select a match?
The area filter plays an important role in reducing computation time, as the number of targets is significantly decreased. Furthermore, since the area filter is still wide, icebergs that are only partially visible or whose contours are enlarged by neighbouring icebergs are not excluded as targets. The solidity filter has a great influence on target reduction, as often only one target remains after filtering. In this case, the remaining target is automatically selected. This has a positive effect on the computational complexity, as the matching step is omitted. However, if the correct target is not visible and only one other target remains, this target is automatically selected as a match, leading to an erroneous result. Moreover, the narrow solidity filter may cause the correct target to be excluded if its solidity value is altered by the presence of adjacent icebergs or sea ice.

To what extent is the algorithm able to recognize icebergs that are only partially visible in a satellite image?
Very large icebergs (>1000 km^2) are often only partially visible in SAR images with a swath width of 250 km. By segmenting the contours before matching, they can still be detected. These very large icebergs are very scarce and therefore often the only target within the area and solidity range, which makes tracking straightforward.

How does the performance of the method compare to an already existing method?
The CC method is compared to the Centroid Distance Histogram (CDH) method, as described by Koo et al., 2021. For small to large icebergs (>10 and <1000 km^2), the performance of both methods is similar.

For small icebergs (<10 km^2), the performance of both methods is poor, this is mainly because icebergs in this area range are abundant, making it hard to select the correct target among other targets. For medium to large icebergs (10 to 1000 km^2), the performance is divergent with one method occasionally outperforming the other. For very large icebergs (>1000 km^2), the CC method with a recall of 0.41 clearly outperforms the CDH method with a recall of 0.10.

6.1. Recommendations

There are several suggestions on how this study can be used in future research. Improvements on this study can be made by introducing a threshold that prevents automatic selection of wrong targets. Secondly, a multiscale method can be introduced that depends on the size of the tracked icebergs, as the extent of the search area is limited by the computational capacity. When tracking medium to very large icebergs, the resolution of the SAR images can be down-scaled and the seed grid spacing can be increased. This reduces the computational complexity and allows for enlargement of the search area. In addition, ocean current data can be used to match the search area with the likely drift direction of the iceberg, thus reducing the probability of the iceberg drifting outside the search area, without increasing the computational complexity.

Furthermore, various combinations of the CC and CDH method can be considered. First, for very small icebergs, the solidity filter can be included in the CDH method so that the number of targets in this area range is reduced more effectively. Second, for medium-sized icebergs, the similarity function of the CDH method can be introduced after the matching step of the CC method as a threshold before selecting a target. In this way, the automatic selection of wrong targets is prevented in the absence of the correct target. Finally, for very large icebergs, the use of Dynamic Time Warp (DTW) can be considered for the similarity calculation of the centroid distance histograms. Very large icebergs are often only partially visible, resulting in a shift in the histogram, while maintaining their overall shape, the use of DTW could allow comparison of the signals by enabling local stretching or compression.

The iceberg detection method using Simple Non-Iterative Clustering (SNIC) is often unable to distinguish the iceberg from its surroundings in specific situations, such as icebergs surrounded by sea ice with high backscatter, or meltwater on the iceberg that reduces the backscatter. To improve iceberg detection, more adaptive approaches could be taken into consideration that have been shown to be more accurate in these circumstances (Barbat et al., 2021).

Bibliography

Bamber, J. L., Oppenheimer, M., Kopp, R. E., Aspinall, W. P., & Cooke, R. M. (2019). Ice sheet contributions to future sea-level rise from structured expert judgment. *Proceedings of the National Academy of Sciences of the United States of America*, *166*, 11195–11200. https://doi.org/10.1073/pnas.1817205116

Barbat, M. M., Rackow, T., Hellmer, H. H., Wesche, C., & Mata, M. M. (2019). Three years of near-coastal antarctic iceberg distribution from a machine learning approach applied to sar imagery. *Journal of Geophysical Research: Oceans*, *124*, 6658–6672. https://doi.org/10.1029/2019JC015205

Barbat, M. M., Rackow, T., Wesche, C., Hellmer, H. H., & Mata, M. M. (2021). Automated iceberg tracking with a machine learning approach applied to sar imagery: A weddell sea case study. *ISPRS Journal of Photogrammetry and Remote Sensing*, *172*, 189–206. https://doi.org/10.1016/j.isprsjprs.2020.12.006

Barbat, M. M., Wesche, C., Werhli, A. V., & Mata, M. M. (2019). An adaptive machine learning approach to improve automatic iceberg detection from sar images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *156*, 247–259. https://doi.org/10.1016/j.isprsjprs.2019.08.015

Braakmann-Folgmann, A., Shepherd, A., Gerrish, L., Izzard, J., & Ridout, A. (2022). Observing the disintegration of the a68a iceberg from space. *Remote Sensing of Environment*, *270*. https://doi.org/10.1016/j.rse.2021.112855

Braakmann-Folgmann, A., Shepherd, A., & Ridout, A. (2021). Tracking changes in the area, thickness, and volume of the thwaites tabular iceberg "b30" using satellite altimetry and imagery. *Cryosphere*, *15*, 3861–3876. https://doi.org/10.5194/tc-15-3861-2021

Budge, J., Vorwaller, S., Cribs, K., & Long, D. G. (2022). The antarctic iceberg tracking database. https://www.scp.byu.edu/data/iceberg/

Budge, J. S., & Long, D. G. (2018). A comprehensive database for antarctic iceberg tracking using scatterometer data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*, 434–442. https://doi.org/10.1109/JSTARS.2017.2784186

Center, U. N. I. (n.d.). About us. https://usicecenter.gov/About

Da Gama Leitão, H. C., & Stolfi, J. (2002). A multiscale method for the reassembly of two-dimensional fragmented objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*, 1239–1251. https://doi.org/10.1109/TPAMI.2002.1033215

Depoorter, M. A., Bamber, J. L., Griggs, J. A., Lenaerts, J. T., Ligtenberg, S. R., Broeke, M. R. V. D., & Moholdt, G. (2013). Calving fluxes and basal melt rates of antarctic ice shelves. *Nature*, *502*, 89–92. https://doi.org/10.1038/nature12567

ESA. (n.d.). Sentinel-1; overview. https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-1-sar/overview

Giorgino, T. (2009). *Computing and visualizing dynamic time warping alignments in r: The dtw package*. http://www.jstatsoft.org/

Google Earth Engine. (n.d.-a). Sentinel-1 sar grd: C-band synthetic aperture radar ground range detected, log scaling. https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD

Google Earth Engine. (n.d.-b). What is earth engine? https://earthengine.google.com/faq/

Koo, Y. H., Xie, H., Ackley, S. F., Mestas-Nuñez, A. M., Macdonald, G. J., & Hyun, C. U. (2021). Semi-automated tracking of iceberg b43 using sentinel-1 sar images via google earth engine. *Cryosphere*, *15*, 4727–4744. https://doi.org/10.5194/tc-15-4727-2021

Liu, H. C., & Srinath, M. D. (1990). Partial shape classification using contour matching in distance transformation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*, 1072–1079. https://doi.org/10.1109/34.61706

Mazur, A., Wåhlin, A., & Krezel, A. (2017). An object-based sar image iceberg detection algorithm applied to the amundsen sea. *Remote Sensing of Environment*, *189*, 67–83. https://doi.org/10.1016/j.rse.2016.11.013

NumPy. (n.d.). *Numpy.gradient* (Version 1.24). https://numpy.org/doc/stable/reference/generated/numpy.gradient.html

Olson, E. (2011). Particle shape factors and their use in image analysis - part 1: Theory. *Journal of GxP Compliance*, *15*.

Pauling, A. G., Bitz, C. M., Smith, I. J., & Langhorne, P. J. (2016). The response of the southern ocean and antarctic sea ice to freshwater from ice shelves in an earth system model. *Journal of Climate*, *29*, 1655–1672. https://doi.org/10.1175/JCLI-D-15-0501.1

Salvador, S., & Chan, P. (2004). *Fastdtw: Toward accurate dynamic time warping in linear time and space*.

Scambos T, S. A., Sergienko O. (2005). Icesat profiles of tabular iceberg margins and iceberg breakup at low latitudes. *Geophysical Research Letters*, *32*, 1–4. https://doi.org/10.1029/2005GL023802

Schodlok, M. P., Hellmer, H. H., Rohardt, G., & Fahrbach, E. (2006). Weddell sea iceberg drift: Five years of observations. *Journal of Geophysical Research: Oceans*, *111*. https://doi.org/10.1029/2004JC002661

Silva, T. A., Bigg, G. R., & Nicholls, K. W. (2006). Contribution of giant icebergs to the southern ocean freshwater flux. *Journal of Geophysical Research: Oceans*, *111*. https://doi.org/10.1029/2004JC002843

Stuart, K. M., & Long, D. G. (2011). Tracking large tabular icebergs using the seawinds ku-band microwave scatterometer. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *58*, 1285–1300. https://doi.org/10.1016/j.dsr2.2010.11.004

Timmermann, R. (2002). Simulations of ice-ocean dynamics in the weddell sea 1. model configuration and validation. *Journal of Geophysical Research*, *107*. https://doi.org/10.1029/2000jc000741

Vernet, M., Geibert, W., Hoppema, M., Brown, P. J., Haas, C., Hellmer, H. H., Jokat, W., Jullion, L., Mazloff, M., Bakker, D. C., Brearley, J. A., Croot, P., Hattermann, T., Hauck, J., Hillenbrand, C. D., Hoppe, C. J., Huhn, O., Koch, B. P., Lechtenfeld, O. J., … Verdy, A. (2019). The weddell gyre, southern ocean: Present knowledge and future challenges. *Reviews of Geophysics*, *57*, 623–708. https://doi.org/10.1029/2018RG000604

Yulmetov, R., Marchenko, A., & Løset, S. (2016). Iceberg and sea ice drift tracking and analysis off north-east greenland. *Ocean Engineering*, *123*, 223–237. https://doi.org/10.1016/j.oceaneng.2016.07.012

Zhan, C., Zhang, L., Zhong, Z., & Didi-Ooi, S. (2018). Deep learning approach in automatic iceberg - ship detection with sar remote sensing data. https://doi.org/https://doi.org/10.1190/IGC2018-437

A

Sensitivity Analysis

For the automatic tracking using the Contour Curvature (CC) method, a lot of settings are defined. To select the best settings and investigate the influence, a sensitivity analysis was performed. In this chapter the chosen settings are mentioned as well as the influence each of these values has on the filtering, segmentation or matching step.

The settings were calibrated using 5 icebergs. For each of the settings in Table A.1 a range of values was tested for each of the icebergs, after which the best performing setting value was selected.

The following settings are maintained in the Contour Curvature method throughout the tracking of the 15 icebergs.

Setting	Value
Peak height	125
Prominence	100
Smoothing factor	1.0
Radius Search Area	125 km
Area range	$[A_{ref}/2.5 \; ; \; A_{ref} \; \text{x} \; 2.5]$
Solidity range	$[S_{ref} - 0.05 \; ; \; S_{ref} + 0.05]$

Table A.1: Settings chosen for the tracking of icebergs throughout the study

The first three values in Table A.1; the peak height, prominence and smoothing factor, are of importance in the contour segmentation step. Each iceberg contour is divided into different segments, based on maxima and minima locations of its curvature function. The peak height describes the minimum (absolute) peak value for a peak to be considered as segmentation position, the prominence describes the minimal peak base width. The smoothing factor s defines the width of the Gaussian filter, which is used to smooth the contour using a convolution. The Gaussian filter width is defined by $1/s^2$.

Figure A.1 illustrates the effect the chosen peak height has on the number of created segments. The prominence value works in a similar way, where a higher prominence factor leads to fewer segments as illustrated in Figure A.2. The effect of the smoothing factor is shown in Figure A.4. When the contour is smoothed more, the curvature peaks essentially gets less high, this is important to take into account if the peak height is already chosen.

The size of the search area is defined by its radius around the centroid of the latest selected match. A greater search radius means more icebergs are included, this leads to greater possibilities of wrong matching. On the other hand, if the search area is too small, the chances of the iceberg drifting outside the search area are higher. For this reason the search area is set to its maximum extent, which is 125 km in this study. A larger search area leads to surpassing of the GEE computational limit.
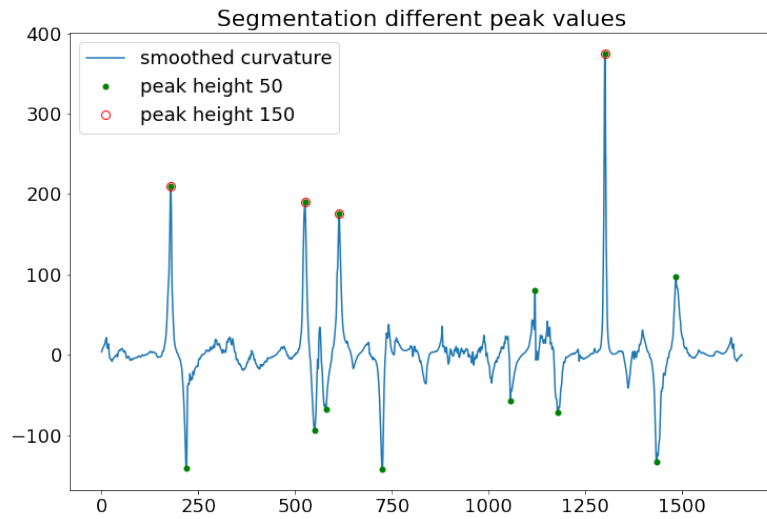
(a)
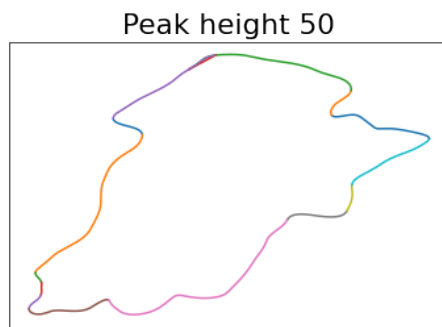
(b)
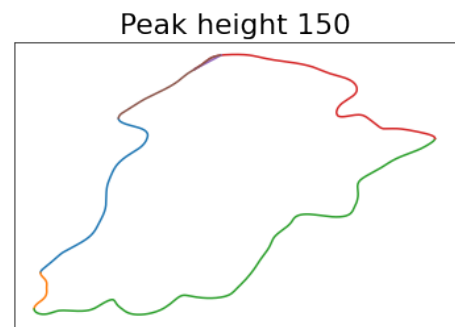
(c)

Figure A.1: a) Curvature function with segmentation positions for a peak height of 50 (green) and 150 (red circle) b) A peak height value of 50 was used, the contour is splitted into 15 segments. c) A peak height value of 150 was used, the contour is splitted into 6 segments. The prominence value was kept constant at 5, the smoothening factor at 1.0.

The area and solidity range define the width of the filters. The area filter is set to include wide range of areas is included, so that partially visible, or icebergs with enlarged areas due to adjacent icebergs are not excluded. The solidity filter is set rather narrow, so that other icebergs with overall different shapes are effectively excluded.

(a)

Prominence 5

(b)

Prominence 250

(c)
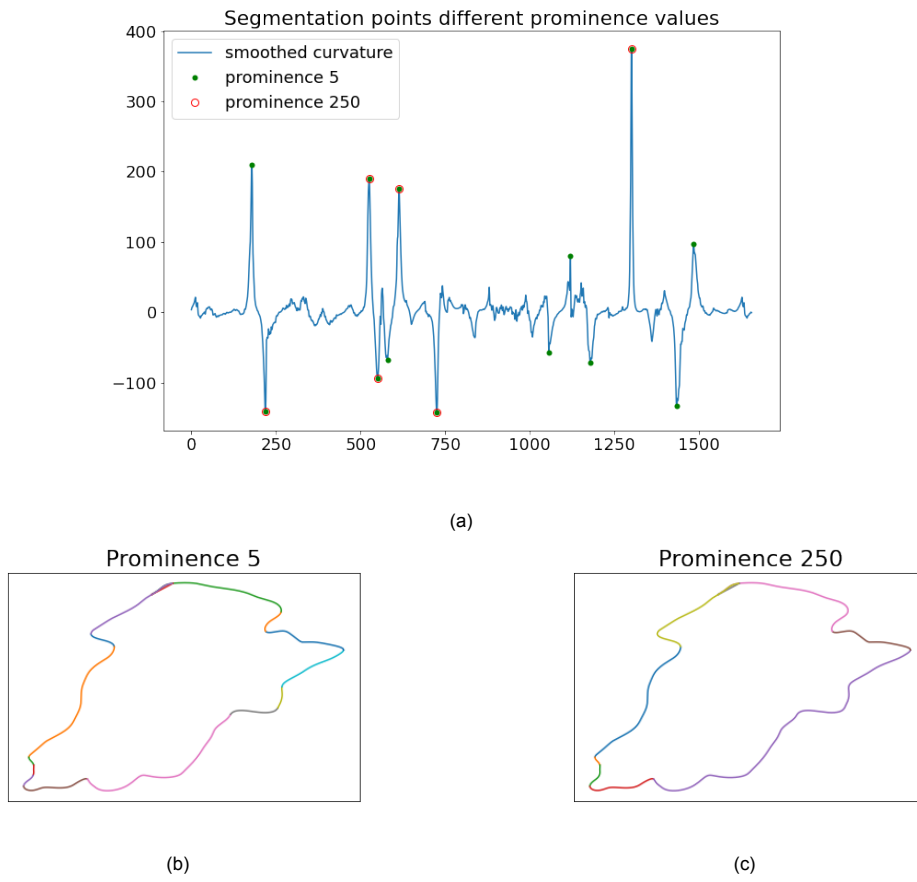
Figure A.2: a) Curvature graph with segmentation positions for a prominence value of 5 (green) and 250 (red circle) b) A prominence value of 5 was used, the contour was splitted into 15 segments. c) A prominence value of 250 was used, the contour was splitted into 9 segments. The peak height value was kept constant at 50, the smoothening factor at 1.0.

(a)
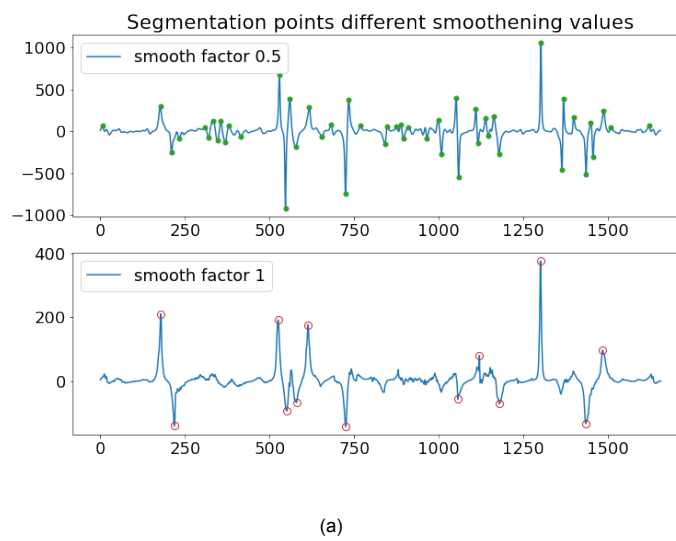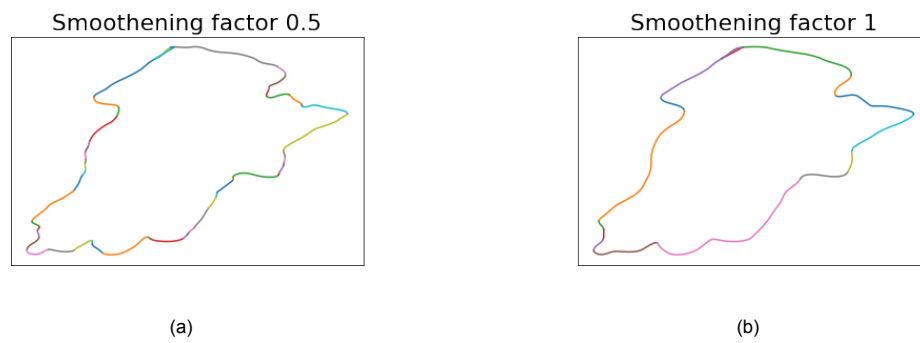
Figure A.4: a) A smoothening factor of 0.5 was used, splitting the contour into 51 segments. Smaller features along the contour are better visible. b) A smoothening factor of 1.0 was used, splitting the contour into 15 segments, less detail is visible in the contour with respect to the 0.5 smoothening factor. The peak height value was kept constant at 50, the prominence value at 5.

B

Overview of Tracked Icebergs

Nr	Image	A [km2]	Sol.	Nr seg.	D [km]	Start date	End date	Start coord.	End coord.	Images
1		239	0,7955	3	3409	2015-10-20	2016-10-20	[-27.20, -74.27]	[-36.92, -75.25]	187
2		539	0,8929	6	5622	2015-04-28	2016-04-29	[8.95, -69.77]	[-32.85, -73.84]	173
3		450	0,8055	12	3773	2015-07-05	2016-07-05	[-5.57, -69.94]	[-40.47, -70.96]	133
4		66	0,9073	6	5309	2015-06-12	2016-06-11	[15.85, -69.34]	[-32.07, -72.52]	102
5		556	0,8800	6	4528	2015-05-29	2016-04-20	[-7.19, -70.29]	[-30.63, -73.71]	152
6		5600	0,8662	110	3532	2017-12-03	2018-11-28	[-62.20, -68.56]	[-60.47, -66.53]	254
7		360	0,9370	4	2783	2016-03-27	2017-03-27	[-31.18, -74.02]	[-52.10, -67.07]	133
8		559	0,8962	27	3345	2015-08-03	2016-07-20	[-17.81, -72.28]	[-42.17, -71.24]	128
9		16	0,8124	4	4054	2015-08-05	2016-08-13	[-18.82, -72.49]	[-47.71, -65.70]	100
10		3	0,9399	2	2488	2016-07-03	2017-07-23	[-29.26, -73.32]	[-47.76, -64.22]	141
11		78	0,8583	3	6931	2015-04-29	2016-04-28	[5.97, -69.81]	[-48.62, -71.16]	144
12		28	0,8960	4	1879	2016-05-26	2016-12-25	[-38.83, -71.08]	[-35.77, -64.73]	58
13		27	0,9191	2	1919	2016-05-20	2017-01-07	[-38.69, -71.63]	[-36.36, -64.89]	58
14		129	0,7541	8	4960	2016-03-11	2017-03-16	[-11.66, -70.85]	[-46.32, -66.83]	119
15		135	0,8758	9	6443	2018-06-23	2019-06-25	[12.51, -69.46]	[-35.86, -77.46]	305

Table B.1: From left to right; The iceberg number, the contour of the iceberg, A = surface area of the iceberg [km²], Sol.=Solidity, Nr seg.= Number of segments in the contour, D = travelled distance [km], Start date = the starting date for tracking, End date = the end date for tracking, Start coord.= the starting coordinates for tracking, End coord.= the ending coordinates for tracking, Images = the number of SAR images within the tracking period where the iceberg is visible. Note: the travelled distance is computed as the displacement of the centroid, for iceberg 6 this value is most likely distorted by the movement of the centroid when the iceberg is partially visible, resulting in a higher travelled distance. The similar is true for other icebergs, although with a smaller effect.
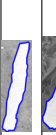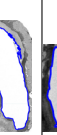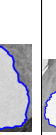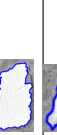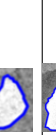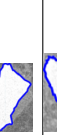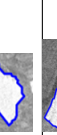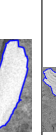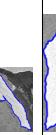
Nr	Image	CC C	CC Recall	CC W	CC NV	CC MF	CC FPR	CC Int	CDH C	CDH Recall	CDH W	CDH Int
1		166	0,888	25	14	11	0,440	3	138	0,738	3	1
2		91	0,526	98	64	34	0,347	0	71	0,410	26	0
3		82	0,617	16	12	4	0,250	1	80	0,602	4	1
4		85	0,833	12	8	4	0,333	8	88	0,863	0	2
5		98	0,645	45	20	25	0,556	0	123	0,809	12	0
6		104	0,409	23	15	8	0,348	2	24	0,098	9	2
7		49	0,368	59	25	34	0,576	0	76	0,571	6	0
8		100	0,781	37	28	9	0,243	1	105	0,820	4	0
9		65	0,650	7	5	2	0,286	3	62	0,620	9	0
10		35	0,248	53	41	12	0,226	4	22	0,156	112	3
11		109	0,757	5	4	1	0,200	0	106	0,736	1	2
12		22	0,379	15	6	9	0,600	2	35	0,603	3	0
13		14	0,241	32	20	12	0,375	2	32	0,552	11	1
14		88	0,739	12	11	1	0,083	0	88	0,739	2	0
15		267	0,875	2	1	1	0,500	2	217	0,711	8	3

Table B.2: CC = Contour Curvature method, CDH = Centroid Distance Histogram method | From left to right; The iceberg number, the contour of the iceberg, C = the number of correctly found matches, Recall = the Recall, W = the number of wrongly matched targets, NV = the amount of wrong matches where the iceberg was not visible, MF = the amount of wrong matches as result of a faulty matching step, CC FPR = False Positive Rate = CC MF/ CC W, Int = the amount of manual interruptions needed