Calving front monitoring at sub-seasonal resolution: a deep learning application to Greenland glaciers

Erik Loebel¹, Mirko Scheinert¹, Martin Horwath¹, Angelika Humbert²⁻³, Julia Sohn²⁻⁴, Konrad Heidler⁵, Charlotte Liebezeit¹, and Xiao Xiang Zhu⁵

¹Technische Universität Dresden, Institut für Planetare Geodäsie, Dresden, Germany
²Alfred-Wegener-Institut Helmholtz Zentrum für Polar- und Meeresforschung, Sektion Glaziologie, Bremerhaven, Germany
³Universität Bremen, Fachbereich Geowissenschaften, Bremen, Germany
⁴IU Internationale Hochschule, Erfurt, Germany
⁵Technische Universität München, Chair for Data Science in Earth Observation, München, Germany

Correspondence: Erik Loebel (erik.loebel@tu-dresden.de)

Abstract. The mass balance of the Greenland ice sheet is strongly influenced by the dynamics of its outlet glaciers. Therefore, it is of paramount importance to accurately and continuously monitor these glaciers, especially the variation of their frontal positions. A temporally comprehensive parameterization of glacier calving is essential to understand dynamic changes and to constrain ice sheet modelling. However, current calving front records are often limited in temporal resolution as they rely on manual delineation, which is laborious and not feasible with the increasing amount of satellite imagery available. In this contribution, we address this problem by applying an automated method to extract calving fronts from optical satellite imagery. The core of this workflow builds on recent advances in the field of deep learning while taking full advantage of multispectral input information. The performance of the method is evaluated using three independent test datasets. We calculate a mean delineation error of 61.2 m, 73.7 m, and 73.5 m, respectively. Eventually, we apply the technique to Landsat-8 imagery. We generate 9243 calving front positions across 23 Greenland outlet glaciers from 2013 to 2021. Resulting time series resolve not only long-term and seasonal signals but also sub-seasonal patterns. We discuss the implications for glaciological studies and present a first application to the analysis of the effect of bedrock topography on calving front variations. Our method and derived results represent an important step towards the development of intelligent processing strategies for glacier monitoring, opening up new possibilities for studying and modelling the dynamics of Greenland outlet glaciers.

1 Introduction

Over the past two decades, the Greenland Ice Sheet has been a major contributor to sea level rise (Horwath et al., 2022). Models suggest, that this imbalance will continue with a warming climate (Goelzer et al., 2020; Edwards et al., 2021; Rückamp et al., 2020). While about half of the ice mass loss is due to increased meltwater runoff, the other half is due to changes of ice discharge to the ocean related to changes of the ice flow dynamics of outlet glaciers (Otosaka et al., 2023). Several mechanisms act as controls and indicators for dynamic glacier changes. In particular, calving and calving front variations have been identified as crucial parameters for investigating the physical mechanisms of Greenland outlet glaciers (Joughin et al., 2008a; Moon and
Joughin, 2008; Benn et al., 2017; Trevers et al., 2019; Cook et al., 2021; Melton et al., 2022). In addition, recent studies have shown that calving front retreat is associated with increased ice discharge (King et al., 2018; Mouginot et al., 2019; King et al., 2020). An accurate representation of calving front behaviour is therefore an important requirement for constraining ice sheet modelling and improving simulations of future mass loss and sea level contribution (Vieli and Nick, 2011; Bondizo et al., 2017; Morlighem et al., 2017, 2019; Greene et al., 2024). Overall, temporally and spatially comprehensive data products of calving front variation are essential for a better understanding and modelling of marine terminating glaciers.

The steady increase in quality and availability of satellite imagery provides new opportunities for a continuous and accurate monitoring of glacier calving front positions. Nevertheless, current data records mostly rely on manual delineation (Schild and Hamilton, 2013; Joughin et al., 2015; ENVEO, 2017; Andersen et al., 2019; King et al., 2020; Goliber et al., 2022; Black and Joughin, 2023). This is a laborious, time-consuming and therefore ineffective process, given the ever-increasing volume of data. Thus, such calving front products often lack temporal resolution, making seasonal analysis and associated modelling efforts difficult. In response to the need for scalable processing strategies, several empirical feature extraction algorithms have been introduced over the last decades, all aiming to provide robust automated calving front extraction (Sohn and Jezek, 1999; Liu and Jezek, 2004; Seale et al., 2011; Rosenau, 2014; Krieger and Floricioiu, 2017; Liu et al., 2021). Yet, most of these methods are either not tested for spatial transferability and large-scale application, or require case-specific modifications.

With the advent of deep learning and big data methods in remote sensing, new opportunities have emerged to solve complex image processing tasks (Zhu et al., 2017). In recent years, a number of case studies have used deep Artificial Neural Networks (ANN) to extract calving front positions. Both optical (Mohajerani et al., 2019) and synthetic aperture radar (SAR) (Zhang et al., 2019; Baumhoer et al., 2019) sensors have been used. Based on these case studies, numerous studies have advanced the ANN architecture (Heidler et al., 2021; Marochov et al., 2021; Periyasamy et al., 2022; Davari et al., 2022a,b; Heidler et al., 2023; Herrmann et al., 2023; Wu et al., 2023), have assessed potential input information (Loebel et al., 2022) and have pursued the multi-sensor capability (Zhang et al., 2021). In addition, dedicated data products have been developed for training and validation (Goliber et al., 2022) as well as benchmarking (Gourmelon et al., 2022) of ANN applications. The results from Cheng et al. (2021) and Zhang et al. (2023), namely the CALFIN and AutoTerm repositories, are currently the only two automatically generated data sets of calving front locations with a Greenland-wide scope.

Building on these achievements, this paper discusses the application and extensive validation of a specially tailored deep learning method for automated calving front extraction using Landsat-8 optical imagery. In doing so, we provide a data product for 23 Greenland outlet glaciers from 2013 to 2021. We compare this data product to the automatically delineated CALFIN and AutoTerm as well as to the manually delineated TermPicks and Black and Joughin (2023) repositories. By exploiting the full multispectral sensor information, our method is able to extract a significant amount of calving fronts that could not be extracted by the other automation methods. By achieving this robust and scalable calving front extraction, we meet the glaciology community requirement for a comprehensive parameterization of glacier calving in Greenland and make important steps towards establishing artificially intelligent processing strategies for glacier monitoring tasks. Overall, we provide the wider cryosphere community with a methodology, a data product and implementation, a comparison to existing products, as well as a discussion of glaciological implications.
Section 2 introduces the data and the applied deep learning method for automated calving front extraction. Section 3 gives an assessment of the accuracy of our method and its spatial transferability. In Section 4 we present our data product, the derived time series and a comparison to existing data repositories. As part of the discussion in Section 5, we provide an application of our results to analyse the interaction between calving front change and bedrock topography. Finally, in Section 6 we draw conclusions and provide an outlook.

2 Data and Methods

The presented processing system extracts glacier calving front shapefiles from multispectral Landsat-8 imagery. In this process, we use satellite imagery as reference data and apply a specialized ANN. This involves a series of processing steps and configurations which are explored in the following section.

2.1 Data source

Our processing system is based on optical Landsat-8 imagery. We use the orthorectified and radiometrically calibrated level 1 data products as provided by the United States Geological Survey (U.S. Geological Survey, 2023). Carrying two scientific instruments, the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS), Landsat-8 provides a particularly wide multispectral coverage. The eleven spectral bands comprise data from visible, near-infrared, short-wave infrared and thermal infrared wavelengths, from 0.435 µm to 1.384 µm. With exception to the panchromatic band and the two thermal bands, which have a spatial resolution of 15 m and 100 m respectively, all other bands have a resolution of 30 m. All available bands except band 8 and band 9 are used as input for our ANN. The 15 m resolution band 8 is excluded due to the otherwise too high computational cost. Band 9 is outside an atmospheric window and is therefore intended for atmospheric observations. The integration of multispectral bands leads to generally more accurate predictions than using conventional single-band inputs only, which has already been shown by Loebel et al. (2022). This is especially true for difficult illumination and ice mélange conditions.

2.2 Reference dataset

We use manually delineated calving front locations as reference data. For model training, we use 698 calving front positions across 19 Greenland glaciers between 2013 and 2019. Glaciers are selected for their broad spatial distribution and diverse morphology as well as for different calving and ocean conditions. A spatial overview of all Greenland glaciers applied in this study is given in Figure 1. As the performance of ANN methods highly depends on training data we pay special attention to cover a diversity of morphological features, terminus with heavy crevassing, different calving and ice mélange conditions as well as varying illumination and cloud situations. To test the model we apply three different testing sets. The TUD testing dataset includes four additional Greenland glaciers, Boydell and Drygalski Glacier at the Antarctic Peninsula, Storbreen Glacier in Svalbard as well as Upsala Glacier in Patagonia. In total, the TUD testing set contains 200 calving front positions across 27 glaciers from 2020 and 2021. In addition to our own testing data set we use manually delineated calving fronts from the
Figure 1. Overview map of the 23 Greenland glaciers used in the TUD reference dataset. Glaciers marked by a red dot are used for training and testing. White dots indicate glaciers only used for model testing. Not on this map: Boydell Glacier (Antarctica), Drygalski Glacier (Antarctica), Storbreen Glacier (Svalbard) and Upsala Glacier (Patagonia) which are only applied for model testing. The basemap is taken from the QGreenland package (Moon et al., 2022).

ESA-CCI (ENVEO, 2017) and the CALFIN (Cheng et al., 2021) product. Here, we use all available calving front positions for our selected Greenland glaciers for which we find a corresponding Landsat-8 scene with less than 24 hours time difference. This results in additional 100 manually delineated calving front positions for the ESA-CCI and 110 for the CALFIN testing datasets.

2.3 Delineating calving fronts by deep learning

For automated calving front extraction, we apply a modified version of the approach published by Loebel et al. (2022). The main difference is, that we only use the multispectral information and no textural and topographic features. This reduces the input from 17 to nine layers. Additionally, we have expanded the reference data set by 170 entries. These new calving front
traces focus specifically on cloudy, low illumination and scene border conditions, thereby enhancing the method in this regard. Figure 2 gives a broad overview of the processing workflow.

2.3.1 Pre-Processing

To use the satellite data as input for the ANN requires pre-processing. In particular, we create stacked raster subsets from the multispectral satellite bands and the manually delineated calving front locations. These subsets have dimensions of 100 px × 512 px with a unified 30 m ground sampling distance and are centered on the calving front of the respective glacier. The 30 m ground sampling distance, and thus the exclusion of panchromatic band 8, is a compromise between the spatial context provided within a single subset, the computational effort and the resolution of the calving front predictions. For each multispectral band we apply an image enhancement in form of a cumulative count cut, clipping the data between the 0.1 and 98 percentile, counteracting overexposure in our satellite imagery. Additionally, all satellite bands are then normalized to the range between 0 and 1 using an 8-bit quantization. Corresponding manually delineated calving front positions, given either as a line string or polygon shape-file, are processed into binary raster masks segmenting land and glacier from ocean. Altogether, one stacked raster subset includes nine satellite bands and a matching ground truth mask.

2.3.2 Semantic image segmentation

To extract the calving front location from the input images we apply a convolutional neural network that performs a pixel-wise semantic image segmentation, separating a glacier-land class from a water class. In particular, we use a U-Net type architecture introduced by Ronneberger et al. (2015). This architecture consists of a contracting path, resembling a typical convolutional network where spatial resolution is reduced while feature information is increased, followed by an expanding path where feature and spatial information are combined. The receptive field of a U-Net is defined by the number of contracting and
expanding blocks. As calving front extraction needs adequate spatial context (Heidler et al., 2021) in this study we enhance the U-Net by two additional resolution levels, i.e. from four to six.

Our model is fitted using the pre-processed training data. Before initializing the model training we select every fifth image of the training dataset for internal validation. The remaining training data is augmented eightfold by rotating and flipping. Finally, the resulting 6208 raster subsets are used for fitting the model. For this, we use randomized batches of size eight and apply the Adam optimization algorithm (Kingma and Ba, 2014) on a binary cross-entropy loss function for a total of 200 epochs. Final model weights are selected based on the classification accuracy of the internal validation dataset.

The ANN processing is implemented using the TensorFlow 2.4 library (Abadi et al., 2015). Model training is carried out on an IBM Power 9 node and an NVIDIA V100 GPU with 32 GB high bandwidth memory. The training of one model requires about twelve hours with a main memory utilization of 80 GB and an average GPU power consumption of 265 W.

2.3.3 Post-Processing

As output of the ANN output we derive a floating point number probability mask where each image pixel is assigned a probability between 0 (water) and 1 (glacier and land). During post-processing, we vectorize this probability mask using the Geospatial Data Abstraction Library (GDAL) contour algorithm (GDAL/OGR contributors, 2020) with a threshold of 0.5 and separate the longest feature. Eventually, we extract the glacier’s calving front by intersecting the vectorized coastline trajectory with a static mask. This mask is created manually for each glacier and specifies a corridor of possible calving front locations. Calving fronts exceeding the 512 px × 512 px window are split into multiple independent predictions which are then averaged in the overlapping area before vectorization. Applying this strategy, which is motivated by Baumhoer et al. (2019), Zachariae Isstrøm, Nioghalvfjerdsbrae and Humboldt Glacier are split into two, three and seven separate overlapping predictions, respectively.

3 Accuracy assessment

Our own TU Dresden (TUD) testing set contains 200 labeled images from the years 2020 and 2021. We emphasize that they are temporally separated from the years of the training datasets. To ensure spatial transferability of our method this test data set includes imagery for additional four Greenland glaciers, two glaciers at the Antarctic Peninsula, one glacier in Svalbard and one glacier in Patagonia. In addition to our own testing dataset we apply another 100 manually picked calving fronts provided by the ESA Greenland Ice Sheet CCI project (ENVEO, 2017) and 110 provided by the CALFIN product (Cheng et al., 2021).

3.1 Error Estimation

The distance between the predicted and the manually delineated calving front is taken as the main error measure in the model testing. We calculate the average minimal distance error by averaging the minimal distances between the predicted front trajectory and the manual delineation calculated every 30 m. Our definition of the average minimal distance error is comparable to the estimates used by Cheng et al. (2021) and Zhang et al. (2023). Figure 3 illustrates some test results for diverse testing
Figure 3. Test results for example scenes from the (a-f) TUD, (g,h) CALFIN and (i-k) ESA-CCI testing set. Manually delineated calving fronts are depicted as dashed black lines. The ANN prediction is shown in orange. The mean average minimal distance error for the respective scene is given both in meters and in pixels. All depicted results derive from the same fitted ANN model. Landsat-8 imagery courtesy of the U.S. Geological Survey.

images from the three test sets. Along with the manually picked calving front (dashed black) and the ANN delineated calving front (orange) the average distance between them is indicated. The ANN model reliably delineates calving front locations under
Figure 4. Accuracy assessment for the three independent test datasets. Every horizontal line inside the ‘violin graphs’ represents one of the 50 trained model applied to the test dataset. The vertical extent of each graph is defined by the corresponding minimum and maximum values.

a wide range of ocean, sea ice and ice mélange situations. Furthermore, the model is also able to handle images affected by challenging cloud (Fig. 3d, j) and illumination (Fig. 3c) conditions, as well as calving fronts near the edge of a satellite scene (Fig. 3e). Test images showing large errors are associated to delineation subjectivity (Fig. 3f,h,i) or even human error (Fig. 3k).

Since the ANN training is stochastic every fitted model performs slightly different on our testing data. To ensure statistical stability for a broader numerical assessment we train and test 50 models using the same reference data and model parameters. In order to assess the distance error, we report both the mean and median over the scenes in the test data set. The test results for these 50 models are shown in Figure 4. Whereas the mean distance error is sensitive towards outliers the median distance error informs about systematic model overfitting and general scene-by-scene performance. Since each of the three testing data sets originated from its own, hence independent, imagery, resulting error estimates are not directly comparable. Nevertheless, we suspect that the lower distance error yielded for the TUD testing set is due to the fact that it was generated by the same people who inferred the training data for these models. Overall, the mean average minimal distance errors are comparable to the results from Cheng et al. (2021) and (Zhang et al., 2023) who estimated who estimated $86.7 \pm 1.4$ m and 79 m, respectively.

Table 1 gives the corresponding statistics.

In addition to the average minimal distance estimate we also calculate the Hausdorff distance (Huttenlocher et al., 1993). The Hausdorff distance only considers the greatest distance of all minimum distances along the two trajectories. As longer fronts are more likely to include misclassified parts, this measure tends to be larger for longer fronts. Goliber et al. (2022) applied the median Hausdorff distance to duplicated delineation from different authors in order to estimate the accuracy level of manual
Table 1. Results of the accuracy assessment. Given are the average minimal distance and the Hausdorff distance for the TUD, ESA-CCI and CALFIN test set. For both estimates we provide mean and median values. The standard deviations result from the 50 different models.

<table>
<thead>
<tr>
<th>Test dataset</th>
<th>Average minimal distance</th>
<th>Hausdorff distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (m)</td>
<td>Median (m)</td>
</tr>
<tr>
<td>TUD</td>
<td>61.2 ± 7.5</td>
<td>28.3 ± 1.4</td>
</tr>
<tr>
<td>ESA-CCI</td>
<td>73.7 ± 2.9</td>
<td>45.9 ± 1.4</td>
</tr>
<tr>
<td>CALFIN</td>
<td>73.5 ± 3.3</td>
<td>43.6 ± 1.6</td>
</tr>
</tbody>
</table>

digitization. Depending on the paired authors this manual delineation error varies between 59 m and 7350 m, with an average of 107 m. The median Hausdorff distances calculated for our test data is therefore within the range of manual delineation errors, but slightly larger than the overall author-to-author error of 107 m calculated by Goliber et al. (2022). Altogether, the quality of calving fronts delineated by our ANN model is comparable to that of manually delineated calving fronts.

3.2 Spatial transferability

In addition to the accuracy assessment over the entire test data set, we evaluate the degree of model generalisation and hence the spatial transferability of our method. Out of our 200 test scenes, 61 scenes are from glaciers that are not included in the training data. For these 61 test scenes over our 50 trained models, we calculate a mean (and median) average minimal distance error of 71.3 ± 19.4 m (median: 24.6 ± 2.1 m). This test error is larger than the error over the entire test set, at 61.2 ± 7.5 m. It is thus also larger than the error over the 139 test scenes of glaciers that are part of the training set, at 56.0 ± 5.3 m (median: 30.3 ± 1.7 m). Notably, we see not only a larger test error, but also a higher standard deviation between the models. This is due to a lower success rate and the resulting high error for individual predictions in cases where the ANN failed to locate the calving front.

Figure 5 gives the test results for four example scenes. Depicted glaciers are outside the training dataset. The calving fronts of Tracy Glacier (Fig. 5a), Upsala Glacier (Fig. 5c) and Drygalski Glacier (Fig. 5d) are reliably extracted with low distances to the manually delineated reference, and low deviation among all trained models. The accuracy is comparable to that of glaciers within the training data set. In contrast, the extractions for Storbreen Glacier (Fig. 5b, left) have a large error and high deviation among the trained models. The calving front is not delineated reliably. This could be due to a combination of the difficult lighting and the snow covered sea ice, which is a condition that might not be adequately represented in the training data. Interestingly, the calving front of the neighboring Hornbreen Glacier (Fig. 5b, bottom right) is extracted accurately over all models.

Among the 61 test images outside the glaciers of the training dataset 57 have an average minimal distance error below 100m (93%), compared to 178 out of 200 over the whole test dataset (89%) and 121 out of 139 of test images of glaciers that are included in the training set (87%). Overall, this assessment confirms the spatial transferability of our processing system.
Figure 5. Test results for example glaciers which are outside the training dataset. Specifically of (a) Tracy Glacier in Greenland, (b) Storebreen in Svalbard, (c) Upsala Glacier in Patagonia and (d) Drygalski Glacier in Antarctica. Orange lines show the predictions from our 50 models. Overlap of lines is indicated by higher color intensity. The average minimal distance metric for each scene is given in meters. Landsat-8 imagery courtesy of the U.S. Geological Survey.

However, the accuracy is lower compared to the extraction from glaciers that were included in the training data. Similar findings have been reported by previous studies (Baumhoer et al., 2019; Cheng et al., 2021; Zhang et al., 2023).

4 Results

4.1 Data product for Greenland from 2013 to 2021

Having trained and tested the ANN model, we apply our processing to Landsat-8 imagery in order to generate temporally dense calving front time series for 23 Greenland outlet glaciers. In doing so, we download Landsat-8 imagery acquired between March 2013 and December 2021. Images with cloud cover larger than 20 % and all Systematic Terrain Correction (L1GT) scenes are manually checked before downloading. Depending on the glacier 51 % (for Ingia Isbræ) to 63 % (for Helheim Glacier) of the available satellite scenes are discarded before download. After ANN processing, failed calving front extractions are discarded. Calving front extraction fails when the predicted coastline trajectory does not intersect the static mask. Finally, calving fronts are filtered using the time series. For this we separate all entries with an area difference of larger than $1 \text{ km}^2$ to both the previous and the next entry. Separated entries are checked manually. Out of the 10587 satellite scenes processed by our ANN, 1344 calving front predictions (13 %) were discarded. Figure 6 gives a tabular overview of the final data product (for locations see Figure 1). In total, we provide 9243 calving front lines, mostly achieving sub-weekly sampling outside polar night. Due to overlapping satellite orbits, glaciers in north, northeast and northwest Greenland undergo up to six image acquisitions per week depending on weather and season. Since we use optical data in this study our time series has observation gaps during polar nights. Depending on latitude, this gap lasts about one month for glaciers in south Greenland and up to three months for glaciers in north Greenland.
Figure 6. Temporal coverage of our ANN generated time series. The numbers and the color intensity indicate the amount of processed calving front positions in the respective year. Glaciers are sorted by latitude from south (left) to north (right).

Figure 7. Rectilinear box method applied to the ANN generated calving front time series for Jakobshavn Isbræ (west Greenland). The glacier, which is separated into a northern and a southern branch, and the calving fronts are shown on the left. The corresponding time series are depicted on the right. Here, calving front positions, expressed as a surface area, are marked by a dot. For the TUD product (black) solid lines connect frontal positions of each year. Time series from the ESA-CCI product (blue) are shown for comparison. Landsat-8 image courtesy of the U.S. Geological Survey.

4.2 Long term, seasonal and subseasonal calving front changes

Marine terminating glaciers experience calving front variations at different time scales. While long-term changes are easy to resolve using already available data products, our time series offers unique opportunities to analyze seasonal and sub-seasonal
terminus changes. To quantify these calving front changes we apply the well-established rectilinear box method (Moon and Joughin, 2008). Rather than using a single profile to measure advance or retreat this method adopts a rectilinear box, thus accounting for uneven changes along the calving front. Figure 7 shows the method applied to our calving front time series for Jakobshavn Isbræ which is separated into a northern and a southern branch. The inferred calving front variation exhibits a pronounced annual pattern combined with smaller sub-seasonal fluctuations. For comparison, the derived time series of the manually delineated ESA-CCI product is shown. Although both datasets agree very well when it comes to comparing singular epochs, the ESA-CCI time series does not reliably capture the temporal variations. This is particularly evident for the year 2014 when a whole annual cycle is missed by the manually delineated product.

Figure 8 presents twelve more examples of our ANN generated time series. Most of these glaciers exhibit pronounced seasonal and sub-seasonal variations overlaid by a long-term signal. Except for Kangiata Nunaata Sermia (Fig. 8a), Ryder Glacier (Fig. 8b) and Hayes Glacier (Fig. 8h), all example glaciers are retreating during the analysed time period. Notably, Zachariae Isstrøm and Humboldt Glacier show an area loss of about 120 km$^2$ and 100 km$^2$ respectively. Ryder Glacier (Fig. 8b) and Nioghalvfjerdsbræ (Fig. 8d) are the only among the 23 glaciers in our study that do not undergo a pronounced seasonality. In those cases, the calving front variation is characterized by a steady advance and the sporadic detachment of large kilometer-sized icebergs. The date of detachment is precisely pinpointed by the time series. In the case of Nioghalvfjerdsbræ (Fig. 8d), the time series also resolves two separate break-offs that occurred in close succession. Other glacier time series, like Hayes Glacier (Fig. 8h), Tracy Glacier (Fig. 8j), Docker Smith Glacier (Fig. 8k) and Harald Moltke Bræ (Fig. 8l), reflect a change in calving rate during our observation period. For Harald Moltke Bræ (Fig. 8l) the onset of this calving front retreat, starting 2019, coincides with the end of its six year-long surging phase and has already been anticipated by Müller et al. (2021).

4.3 Comparison to existing data products

In addition to the data set produced in this study, there are two other automatic delineation products with a circum-Greenland coverage: the CALFIN data set by Cheng et al. (2021) and the AutoTerm repository by Zhang et al. (2023). Additionally, there are a number of manually picked data records. Two particularly comprehensive databases are the TermPicks product (Goliber et al., 2022) and the dataset of Black and Joughin (2023). TermPicks (Goliber et al., 2022) is a compilation of manually delineated calving front data from 19 different authors across 278 glaciers. The Black and Joughin (2023) dataset was created to study weekly and monthly calving front variability. For this purpose, the authors digitised 199 glaciers with a monthly frequency and 20 glaciers with a six-day frequency over a seven-year period. In this section, we will compare these four "big data" Greenland calving front datasets with the results of this study. The comparison takes place on three levels. Firstly, we compare the general statistics and scope. Secondly, we compare results over a reference period and reference glaciers defined according to their temporal and spatial overlap. Thirdly, we examine individual examples.

Table 2 (columns 2 to 4) presents the general statistics for the four datasets. Our data set covers a relatively short time span since we process imagery from the OLI and TIRS Landsat sensors, which have only been available since 2013. With 9243 mapped calving fronts over 23 glaciers our data product is smaller in both scope and size than the CALFIN, AutoTerm, TermPicks and Black and Joughin (2023) products. When examining the number of calving front traces, it is important to
understand that the definition of what a single calving front contains varies from study to study. For instance, a single data entry in our dataset for the Upernavik Isstrøm includes four calving front features. CALFIN lists three separate calving fronts for the same glacier, and AutoTerm and Termpicks list two. For Jakobshavn Isbræ, CALFIN considers the north and south
Table 2. Comparison of the CALFIN, AutoTerm, Black and Joughin (2023), TermPicks product as well as the data product presented in this study. The reference period (2015 to 2019) and the reference glaciers (13 glaciers) are defined by the temporal and spatial overlap of the four data products.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Glaciers</th>
<th>Mapped fronts</th>
<th>Time span</th>
<th>Reference period and glaciers</th>
<th>Mapped fronts</th>
<th>Sampling rate (yr⁻¹)</th>
<th>Unique entries</th>
</tr>
</thead>
<tbody>
<tr>
<td>This study (Loebel et al., 2023)</td>
<td>23</td>
<td>9243</td>
<td>2013-2021</td>
<td>2423</td>
<td>37.28</td>
<td>281</td>
<td></td>
</tr>
<tr>
<td>CALFIN (Cheng et al., 2021)</td>
<td>66</td>
<td>22678</td>
<td>1972-2019</td>
<td>956</td>
<td>14.71</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>AutoTerm (Zhang et al., 2023)</td>
<td>295</td>
<td>278239</td>
<td>1984-2021</td>
<td>6512</td>
<td>100.18</td>
<td>2545</td>
<td></td>
</tr>
<tr>
<td>Black and Joughin (2023)</td>
<td>219</td>
<td>23333</td>
<td>2015-2021</td>
<td>2187</td>
<td>33.65</td>
<td>676</td>
<td></td>
</tr>
<tr>
<td>TermPicks (Goliber et al., 2022)</td>
<td>278</td>
<td>39060</td>
<td>1916-2020</td>
<td>1806</td>
<td>27.78</td>
<td>271</td>
<td></td>
</tr>
</tbody>
</table>

branch separately, while in our data set they are counted as one calving front. In addition, some of our predictions also include smaller neighboring glaciers that are located on the same image tile (e.g. Farquhar Glacier which is included together with Tracy Glacier or Akullersuup Sermia which is included together with Kangiata Nunaata Sermia). Usually, this applies to glaciers of a single glacier system that were previously connected. When counting shape file features, the number of entries in our data product contains 15150 entries.

To better compare the data products and differences in processing strategy, we define a reference period and reference glaciers by considering the temporal (2013 to 2019) and spatial overlap (13 glaciers) of the four data sets. These glaciers are: Kangiata Nunaata Sermia, Helheim Glacier, Kangerdlugssuaq Glacier, Jakobshavn Isbræ, Sermeq Avangnardleq, Store Glacier, Rink Isbrae, Ingia Isbrae, Upernavik Isstrøm, Hayes Glacier, Sverdrup Glacier, Kong Oscar Glacier and Døcker Smith Glacier. Within this reference we looked at mapped fronts, sampling rate and unique entries. Results are given in Table 2 (columns 5 to 7). We consider only one calving front entry per day and per glacier. Effectively, this removes (1) duplicate delineations of the same scene (e.g. from multiple authors in the TermPicks data base or from the reference data which is included in CALFIN data set) and (2) inconsistencies in what constitutes a single calving front entry. Although having the same Landsat data basis our data product achieves a higher sampling rate and more unique front extractions than CALFIN. This is likely due to differences in input feature selection and processing. AutoTerm has the most mapped and unique fronts as well as the highest sampling rate. This is mainly due to the ability to process multi-sensor imagery and the resulting larger database, which includes Landsat, Sentinel-2 and Sentinel-1. A clear advantage over our approach, which is limited to the use of multispectral Landsat data. The manually picked TermPicks and Black and Joughin (2023) data sets have a sampling rate that is lower, but still comparable, to that of our product. It should also be noted that the sampling rate of all five calving front products, and thus the number of mapped fronts, is unevenly distributed across glaciers. This is due to varying satellite image availability and quality but, specifically for manual digitalized products, also due to time constraints and prioritisation. This is evident in the TermPicks database, which has a significantly higher sampling in West than in East Greenland (Goliber et al., 2022) but also for the Black and Joughin (2023) data set, where eight of the 13 glaciers in our reference have a six-day sampling,
Figure 9. Comparison of the CALFIN, AutoTerm, Black and Joughin (2023), TermPicks product (blue) to the data product presented in this study (black) for four example glaciers. Time series are derived along the central flow line of the glacier. Every comparison specifies the mean distance \( d \) between calving front delineations at identical days. While the remaining glaciers have monthly sampling. Overall, our method achieves the second highest sampling rate within this reference and with that 281 out of the extracted 2423 calving fronts were not extracted by CALFIN, AutoTerm, Black and Joughin (2023) or TermPicks.

Figure 9 shows the time series of CALFIN, AutoTerm, Black and Joughin (2023) and TermPicks compared to our study for four individual glaciers. To maximise the sampling of the different data sets, we analyse a centre line profile instead of using the box method. The mean distance for same-day calving front acquisitions \( d \) is indicated for each pair of time series. When examining these four examples, we observe a generally good agreement between the time series. Significant differences exist
only for Humboldt Glacier (Fig. 9d). Here, the data quality of the AutoTerm product seems to be notably worse than for the other glaciers, with large fluctuations up to 5 km in distance. This may be attributed to the glacier front’s large size, which, at least in our method, required additional processing steps. For Kangiata Nunaata Sermia (Fig. 9b) our data product is the only one which captures the signal from the seasonal ice tongue (Motyka et al., 2017; Moyer et al., 2017). This is reflected on the one hand in gaps in the other data sets and on the other hand in a higher distance for some same-day acquisitions. Although Landsat imagery is available, both CALFIN and AutoTerm have almost no calving front traces during the emergence, presence and disintegration of this seasonal ice tongue. We suspect, that the multispectral input information of our processing leads to a better extraction rate for scenes under these challenging conditions. All four examples highlight the varying sampling rates of the data products. In particular, the AutoTerm and Black and Joughin (2023) dataset, which use Sentinel-1 SAR imagery, have coverage during polar night (see late 2017 in Fig. 9a and late 2016 in Fig. 9c). The sampling rate of the TermPicks repository is lower than that of the automated processing systems in these four examples.

Compared to CALFIN, AutoTerm, Black and Joughin (2023) and TermPicks, our data product has the lowest coverage and smallest amount of overall mapped calving front traces. However, due to different processing and the addition of multispectral input information our method is able to extract a significant amount of calving fronts, 12% within the reference, that could not be extracted by the other methods. Importantly, these 12% are likely to include extractions under challenging ice mélange and illumination conditions. For the analyzed reference, our method has a temporal resolution second only to that of the AutoTerm product, which benefits from multi-sensor input imagery. Overall, this comparison also presents a clear argument for the benefits of having multiple data products on glacier calving fronts. Current data products differ in scope but also differ for duplicate extractions for identical glacier front traces, often exceeding estimated delineation uncertainties. A better understanding of these differences is crucial and requires further investigation. As a final point, we want to emphasise the potential of combining different glacier front products (Goliber et al., 2022). Particularly for data sets based on optical data, this not only increases the overall sampling rate, but also allows data gaps to be filled during the polar winter. Greene et al. (2024) have demonstrated the advantages of such a combination for large-scale glaciological analyses.

5 Discussion

Changes in calving front position are, along with other observables like ice velocity and elevation change, part of a complex feedback cycle between a glacier and its environment. Long-term calving front trends of Greenland glaciers are well characterised (Howat and Eddy, 2011; King et al., 2020; Fahrner et al., 2021; Black and Joughin, 2022; Greene et al., 2024). However, about 80% of Greenland glaciers also experience terminus changes on a seasonal and subseasonal basis (Black and Joughin, 2023). A visual inspection of the time series shows that 19 of the 23 glaciers analyzed in this study show a seasonal pattern between the years 2013 and 2021. As observed by other studies (Joughin et al., 2008b; Seale et al., 2011; Carr et al., 2013; Schild and Hamilton, 2013; Murray et al., 2015; Moon et al., 2015; Cassotto et al., 2015; Kehrl et al., 2017; Fried et al., 2018; Sakakibara and Sugiyama, 2020; Kneib-Walter et al., 2021; Black and Joughin, 2023), glacier retreat typically starts in late spring with retreat rates peaking in late summer. A number of mechanisms have been identified as controls for these seasonal
terminus changes. These include the duration and timing of meltwater runoff (Sohn et al., 1998; Nick et al., 2010; Chauche et al., 2014; Carroll et al., 2016; Fried et al., 2018; Wood et al., 2021), changes in buttressing force by sea ice and ice mélange (Howat et al., 2010; Carr et al., 2013; Todd and Christoffersen, 2014; Cassotto et al., 2015; Moon et al., 2015; Kehrl et al., 2017; Robel, 2017; Kneib-Walter et al., 2021), basal sliding (De Juan et al., 2010; Moon et al., 2015) and ocean driven melt (Motyka et al., 2003; Bevan et al., 2012a; Chauche et al., 2014; Carroll et al., 2016). When a glacier is forced into a state of retreat, both the rate and pattern of retreat are modulated by the subglacial topography. For marine terminating glaciers in Greenland, this effect has been studied (Warren, 1991; Warren and Glasser, 1992; Joughin et al., 2008b; Carr et al., 2015; Lüthi et al., 2016; Kehrl et al., 2017; Bunce et al., 2018; Catania et al., 2018; Felikson et al., 2021) and modelled (Enderlin et al., 2013; Morlighem et al., 2016; Choi et al., 2017) intensively. In particular, faster retreat rates were found to be associated with overdeepening and retrograde topography.

The new generation of automatically delineated calving front data products not only facilitates glaciological analysis in terms of significant time savings, but may also provide new insights due to the high temporal resolution and spatial coverage. Figure 10 shows our calving front time series for three example glaciers in relation to bedrock elevation, taken from the Bed-Machine Greenland Version 5 model (Morlighem, 2022). Profiles extend from point A to point B along a central flowline. The calving front of Ingia Isbræ (Fig. 10a) retreated from 2013 to the end of 2017 by 3.2 km (8.6 km² in area) with a pronounced seasonal pattern. Due to retrograde topography (at ~4 km in Fig. 10a), this retreat is particularly rapid in 2016 and 2017. Since 2018, the calving front has been in a topographic minimum, preventing further retreat and reducing seasonal amplitude. These observations confirm the analysis of Catania et al. (2018), which described the continuous retreat of Ingia Isbræ from 2002 to 2016 and suggested further retreat by more than 1 km until the calving front stabilises on the prograde bed topography. The calving front change of Kangerdlugssuaq Glacier (Fig. 10b) shows high seasonal amplitudes as well as a significant retreat from 2016 to early 2018. With the exception of 2017 and 2018, where we observe a sustained retreat, the seasonal amplitude remains almost constant at around 4 km (21 km² in area). This calving front pattern is also described by Kehrl et al. (2017). Furthermore, the authors show that Kangerdlugssuaq’s grounding line has retreated in 2010 and 2011 to a stable bedrock position (bedrock bump at ~10 km in Fig. 10b), resulting in a floating ice tongue of ~5 km in length. Due to the retrograde bedrock topography after the bedrock bump (from 10 km to 15 km in Fig. 10b) and the reinitialization of the seasonal terminus pattern from 2019 to 2021, we suspect that the calving front retreat from 2016 to 2018 is coupled with a grounding line retreat on retrograde bed topography followed by a restabilisation further inland (likely at 15.5 km in Fig. 10b). This would confirm the second scenario suggested by Brough et al. (2019). The calving behavior of Daugaard Jensen Glacier (Fig. 10c) is influenced by the abrupt change of bedrock slope close to the frontal position. An advance beyond this point, from a slightly retrograde into a steeply prograde topography, leads to a loss of basal drag and longitudinal stresses. This influences calving behaviour, and in particular favors calving of tabular icebergs like in 2013 (8.2 km² in size) and 2020 (4.2 km² in size). Although Daugaard Jensen’s glacier front has remained roughly at this location for over 70 years (Stearns et al., 2005) and is considered to be stable (Bevan et al., 2012b), temporally high resolution calving front information is still necessary to resolve and understand these stable glacier dynamics. More generally, high temporal resolution calving front information not only allow to analyze glacier retreat and advance, but also to better differentiate between different calving styles and patterns.
Figure 10. Effect of bedrock topography of calving front variation for (a) Ingia Isbræ, (b) Kangerdlugssuaq Glacier and (c) Daugaard Jensen Glacier. Shown are (from left to right) a satellite image with calving front trajectories as well as a marked profile, bedrock topography and color-coded calving front positions along this profile and the corresponding time series of calving front variation. Note that the axes are scaled differently for each glacier. Landsat-8 imagery courtesy of the U.S. Geological Survey.

6 Conclusions

This study presents a deep learning based processing system for automatic delineation of calving front locations from multi-spectral Landsat-8 imagery. Using three independent test datasets we validate the performance and spatial transferability of our processing system. The quality of the automatically extracted calving fronts is comparable to that of manually delineated calving fronts. Our method enables a considerably higher extraction rate compared to other automation methods that use the same Landsat data basis. The resulting data product, which includes 9242 calving fronts over 23 Greenland glaciers, is therefore a valuable contribution to the existing data repositories.

The presented method and the resulting data product address the needs of the glaciology community for a comprehensive parameterization of glacier calving in Greenland. The time series derived by this processing system resolve long-term, seasonal and sub-seasonal calving front variations. The benefit is particularly significant for large glaciers where there is a lack of manual
delineated data, such as Humboldt Glacier, Zachariae Isstrøm and Nioghalvfjerdsbrae. Due to the spatial transferability of this method, our processing system has the potential to be applied to other marine-terminating glaciers around the world.

By the time series presented in this paper, we give only a selected glimpse into the dynamics of these glaciers. However, the demonstrated capability of automatically resolving the sub-seasonal calving front variations is an important step forward towards a spatially comprehensive Greenland-wide monitoring system. In conjunction with other components concerning ice flow, elevation change, solid earth response and hydrological processes, this will open up new opportunities to integratively assess, model and simulate dynamic ice sheet changes. Advancing towards this digital twin of the Greenland Ice Sheet will improve our understanding of its evolution and its role within the broader Earth climate system.

Intelligent processing strategies, like deep ANN, will play a major role in shaping the future of glacier monitoring and associated modelling tasks. This is especially true for analyzing the increasing amount of remote sensing imagery. Well-trained and thoroughly validated ANN will be state-of-the-art for automated calving front delineation. The results presented in this paper will contribute to future advancements in this field.

**Code and data availability.** The following assets are published along with this article:

- The data product of automatically delineated calving front positions (format: ESRI shapefile), containing 9243 calving front positions across 23 Greenland outlet glaciers, is available at http://dx.doi.org/10.25532/OPARA-208 (Loebel et al., 2023).
- All reference data applied in this study is available at http://dx.doi.org/10.25532/OPARA-282 (Loebel et al., 2024). In particular, this includes 898 manually delineated calving front positions provided in a georeferenced shapefile format, as well as 1220 machine learning ready raster subsets (pre-processed, 9 channels) with their corresponding manual delineated segmentation mask.
- We provide a containerized implementation (platform: Docker) of the presented processing system. The software automatically extracts calving front positions from Landsat-8 or Landsat-9 Level-1 data archives for glaciers used within this study or at user-defined coordinates. This enables the analysis of glaciers that are outside our reference dataset or beyond the temporal frame of our study. The software is available at https://github.com/eloebel/glacier-front-extraction (last access 24 March 2023) and https://doi.org/10.5281/zenodo.7755774 (Loebel, 2023a).
- Our implementation (software: Python 3) of the rectilinear box method is available at https://github.com/eloebel/rectilinear-box-method (last access 24 March 2023) and https://doi.org/10.5281/zenodo.7738605 (Loebel, 2023b).

**Author contributions.** Contributions are according to CRediT. Conceptualization: EL, MS and AH. Data curation: EL and CL. Formal analysis: EL. Funding acquisition: MS, MH, AH, and XXZ. Investigation: EL. Methodology: EL and KH. Software: EL. Supervision: MH and MS. Validation: EL. Visualization: EL. Writing (original draft): EL. Writing (review and editing): EL, MH, MS, KH, AH, CL, XXZ, JS.

**Competing interests.** The author declares that there are no competing interests.
Acknowledgements. We would like to thank the USGS for providing Landsat imagery. Also, we are grateful to the TU Dresden computing centre (ZIH) for providing their high-performance computing and storage infrastructure. We acknowledge the National Snow and Ice Data Center QGreenland package. This work was supported by the Helmholtz Association of German Research Centers as part of the Helmholtz Information and Data Science Incubator, project "Artificial Intelligence for Cold Regions" (AI-CORE, grant no. ZT-I-0016), by the German Federal Ministry of Education and Research (BMBF), project "Greenland Ice Sheet/Ocean Interaction" (GROCE2, grant no. 03F0778G) and by the German Federal Ministry for Economic Affairs and Climate Action in the framework of the "national center of excellence ML4Earth" (grant no. 50EE2201C).
References


Joughin, I., Moon, T., and Black, T.: MEaSUREs Annual Greenland Outlet Glacier Terminus Positions from SAR Mosaics, Version 1, Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center, https://doi.org/10.5067/DC0MLBOCL3EL, (last access: 11.November 2021), 2015.


