



**University of
Zurich** ^{UZH}

Refining Debris-Covered Glacier Outlines Using Land Surface Temperature Data

GEO 511 Master's Thesis

Author

Lorena Müller
18-870-170

Supervised by

Dr. Kathrin Naegeli
Gabriele Bramati
Dr. Hendrik Wulf

Faculty representative

Prof. Dr. Livia Piermattei

20.12.2024

Department of Geography, University of Zurich

Abstract

Accurate mapping of debris-covered glaciers (DCGs) is essential for understanding their dynamics, contribution to sea-level rise and role in regional hydrology. Existing glacier mapping efforts suffer from inaccuracies at the local scale, especially for DCGs. Current approaches often fail to distinguish debris from surrounding terrain due to spectral similarities. In many cases, manual delineation is required, which is time consuming and not suitable for large scale mapping. This thesis developed a methodology to refine DCG outlines using land surface temperature (LST) data, based on the assumption that glacial debris shows a cooler thermal signature influenced by the underlying ice. The Approach Integrates thermal and optical data from Landsat 5 and 8, topographic information from three digital elevation models (Copernicus DEM, NASADEM, and SwissAlti3D), and outlines of existing glacier inventories, including the Swiss Glacier Inventories (SGI2010 and 2016) and the Randolph Glacier Inventory version 7 (RGI 7.0).

A novel index combining LST and near infrared (NIR) data was integrated with topographic information and spectral indices in a random forest classifier. The transferability and robustness of the classifier across different regions was enhanced by using multiple combinations of input variables, which were subsequently integrated into a final classification. The method was calibrated on Zmuttgletscher in the Swiss Alps and extended to Oberaletschgletscher and Unteraargletscher. Temporal analyses for 2003, 2010 and 2016 assessed its robustness over time, while additional applications on Belvedere Glacier (Italy) and Satopanth Glacier (India, Himalayas) evaluated its adaptability to different climatic regions. Statistical comparison with the SGI and RGI inventories yielded validation accuracies between 0.743 and 0.929, often exceeding 0.85. While challenges remain in terms of accuracy and robustness, the method effectively delineates DCG outlines across different glaciers and time periods, providing a scalable and automated approach to DCG mapping. The automated nature of the approach and the use of open-source data make the method a valuable tool for DCG mapping on a global scale.

With upcoming satellite missions such as TRISHNA, SBG and LSTM expected to provide higher resolution thermal data, the methodology is poised for further refinement and wider application. This research advances the use of LST in glaciology and provides a practical tool for improving global glacier inventories.

Contents

Abstract.....	2
Abbreviations.....	6
List of Figures	7
List of Tables.....	9
1 Introduction.....	10
1.1 Motivation and state of research	10
1.2 Research objectives and questions.....	11
1.3 Research aims and structure	12
2 Theoretical Background.....	13
2.1 Thermal infrared imagery and advancing LST retrieval.....	13
2.2 The use of thermal data in DCG mapping	14
2.3 Influence of debris thickness on LST.....	15
2.4 Additional bands used in DCG mapping.....	16
3 Study Sites.....	17
3.1 Zmuttgletscher	17
3.2 Unteraargletscher	18
3.3 Oberaletschgletscher	19
3.4 Belvedere Glacier	19
3.5 Satopanth Glacier	20
4 Data	21
4.1 Satellite Imagery	21
4.1.1 Landsat 5	21
4.1.2 Landsat 8	22
4.2 Digital Elevation Models (DEM)	22
4.2.1 SwissAlti3D	22
4.2.1 NASADEM	23
4.2.2 Copernicus DEM.....	23
4.2.3 SwissAlti3D	23
4.3 Glacier Inventories.....	23
4.3.2 Randolph Glacier Inventory version 7.0 (RGI 7.0).....	24
4.3.2 Swiss Glacier Inventory 2010 (SGI2010).....	24
4.3.1 Swiss Glacier Inventory 2016 (SGI2016)	25
5 Methods.....	26
5.1 Developing the methodology	26

5.1.1 Manual image selection	26
5.1.2 Pixel-wise selection of scene and composite creation.....	27
5.1.3 Normalisation of LST for elevation	28
5.1.4 LST-NIR Index.....	29
5.1.5 Classification with Random Forest.....	30
5.1.6 Statistical metrics for performance evaluation.....	36
5.1.7 Ground Reference.....	36
5.2 Evaluation strategy.....	37
5.2.1 Temporal change analysis.....	37
5.2.2 Cross-glacier analysis	37
6 Results.....	38
6.1 Temporal change analysis	39
6.2 Cross-glacier analysis.....	44
6.2.1 Approach I	44
6.2.2 Approach II	47
6.2.3 Approach III.....	50
6.3 Performance of the LST NIR Index.....	53
7 Discussion	55
7.1 Different Approaches.....	55
7.1.1 Temporal change analysis.....	55
7.1.2 Cross-glacier analysis	56
7.2 Case studies of model performance	59
7.2.1 Distinct value transitions on the main DCG body.....	59
7.2.2 Potential DCG misclassifications and 'noise'	60
7.2.3 Glacier terminus and wideness: the need for accurate thresholding.....	62
7.2.4 Holes in the main DCG body: the case of Belvedere glacier	65
7.2.5 The issue with temporal change analysis on Zmuttgletscher 2010.....	66
7.3 Benefit of LST NIR Index	67
7.4 Limitations.....	67
7.4.1 Coarse resolution and mixed pixels.....	67
7.4.1 Ground reference.....	68
7.4.2 Debris thickness.....	69
7.4.3 Other influences on LST	70
8 Conclusion and Outlook.....	72
8.1 Conclusion.....	72
8.2 Outlook.....	73

Acknowledgements.....	75
References.....	76
Appendix	76
A Google Earth Engine Scripts.....	83
B Classification Results.....	84
C Input Layers.....	89
D Ground reference for 2003 and 2010.....	90
Personal Declaration	92

Abbreviations

ASTER	Advanced Spaceborne Thermal Emission and Reflection
CGI	Comitato Glaciologico Italiano (Italian Glaciological Committee)
COPDEM	Copernicus Digital Elevation Model
DCG	Debris-Covered Glacier
DEM	Digital Elevation Model
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center)
EOS	Earth Observing Systems
ESA	European Space Agency
GED	Global Emissivity Dataset
GEE	Google Earth Engine
LiDAR	Light Detection and Ranging
LST	Land Surface Temperature
LSTM	Land Surface Temperature Monitoring
MSG	Meteosat Second Generation
NASA	National Space Agency
NASADEM	National Space Agency Digital Elevation Model
NDRI	Normalised Difference Rock Index
NDSI	Normalised Difference Snow Index
NDVI	Normalised Difference Vegetation Index
NDWI	Normalised Difference Water Index
NIR	Near Infrared
OBIA	Object-Based Image Analysis
OLI	Operational Land Imager
PGD	Periglacial debris
RF	Random Forest
RGI	Randolph Glacier Inventory
SBG	Surface Biology and Geology
SGD	Supraglacial Debris
SIG	Swiss Glacier Inventory
SNIC	Simple Non-Iterative Clustering
SRTM	Shuttle Radar Topography Mission
SWIR	Shortwave Infrared
TIR	Thermal Infrared
TIRS	Thermal Infrared Sensor
TPI	Topographic Position Index
TRISHNA	Thermal infraRed Imaging Satellite for High-resolution Natural resource Assessment
USGS	United States Geological Survey
VNIR	Visible and Near Infrared

List of Figures

Figure 1: Illustration of a schematic cross-section of a DCG based on Shukla et al. (2010), and the retrieval of LST to distinguish between PGD and SGD. 14

Figure 2: Overview of Zmuttgletscher as the primary study site and its location in the Swiss Alps, showing high-resolution Google satellite imagery overlaid with the SGI2016 glacier outlines..... 18

Figure 3: Overview of the study site of Unteraargletscher and its location in the Swiss Alps..... 18

Figure 4: Overview of the study site of Oberaargletscher and its location in the Swiss Alps. Since the high-resolution Google Satellite Image of this area was snow covered, a scene of Landsat 5 (2010) was used for visualisation..... 19

Figure 5: Overview of the study site of Belvedere glacier and its location in the Italian Alps..... 20

Figure 6: Overview of the study site of Satopanth Glacier and its location in northern India, Himalaya... 20

Figure 7: Flowchart of the process applied in GEE, starting with data selection of relevant input data for the generation of input layers that were used for DCG classification. 26

Figure 8: Scatterplot of the relationship between LST and elevation, exemplary for Zmuttgletscher 2016, showing 3000 points and including a trendline. 28

Figure 9: Normalisation of LST for elevation by regression, visualized for Zmuttgletscher 2016..... 29

Figure 10: Visualisation of the comparison of combining different bands with LST to enhance differentiation between PGD and SGD, using equation (3) while substituting NIR with different bands ('Band x'). 30

Figure 11: Overview of validation accuracies of classification results generated with different approaches, for all glaciers and years tested. 38

Figure 12: Accuracies of multiple band combinations used for the final classification, for the test year, when trained with the same glacier of two other years. The combinations are sorted by mean validation accuracy. The x-axis was adjusted to highlight differences in accuracy. 40

Figure 13: Comparison of validation accuracies per glacier and year with temporal change analysis, creating the final classification mask either with Otsu's threshold or with a manually selected, fixed threshold. The y-axis was adjusted to highlight differences in accuracy. 41

Figure 14: Classification results of Temporal changes, Zmuttgletscher. The colorbar in the final classification sum shows the number of times a pixel was classified as DCG by the different band combinations. 43

Figure 15: Classification results of Cross-glacier analysis, Approach I, Zmuttgletscher. The colorbar in the final classification sum shows the number of times a pixel was classified as DCG by the different band combinations..... 45

Figure 16: RGB image (a) and normalised LST (b) of the selected scene from 11th of October 1998 for the analysis of Satopanth glacier. 47

Figure 17: Classification results of Cross-glacier analysis, Approach II, Zmuttgletscher. The colorbar in the final classification sum shows the number of times a pixel was classified as DCG by the different band combinations..... 49

Figure 18: Classification results of Cross-glacier analysis, Approach III, Zmuttgletscher. The colorbar in the final classification sum shows the number of band combinations. 51

Figure 19: Scatterplot of the relation of LST and NIR for DCG and non-DCG (coloured according to SGI2016 debris-cover), with data from Zmuttgletscher 2016. Visualized in Python..... 54

Figure 20: Visualisation of slope based on the SwissAlti3D (2019) for Zmuttgletscher (a) and Unteraargletscher (b), at a resolution of 2m, used for training and testing of classifier on the respective glaciers in 2016.....	56
Figure 21: Classification sum of Zmuttgletscher 2016 generated with Approach I and overlaid on high-resolution google satellite imagery (a) and zoom to glacier front (b and c). Compared to the SGI2016 glacier outlines.	59
Figure 22: Classification sum of Oberaletschgletscher 2016 generated with Approach III and overlaid on high-resolution google satellite imagery (a), with zoom to glacier front (b and c). Compared to the SGI2016 debriscover outlines.....	60
Figure 23: DCG classification result of Oberaletschgletscher 2010 overlaid on the corresponding Landsat scene and compared to SGI2010 outlines.	61
Figure 24: Classification result for Satopanth glacier with Approach I, showing the DCG outlines of the classifier compared to RGI 7.0 outlines and DCG 'noise' that was filtered out by applying a smoothing filter (a), with zoom to areas on upper south-western glacier arm (b and c).....	62
Figure 25: Final classification sum of Zmuttgletscher 2003, with Approach I (a), and derived outlines based on different thresholds (b). Overlaid on high-resolution Google satellite imagery.....	63
Figure 26: Different outlines for Zmuttgletscher in 2003, generated using Approach I, showing variations based on different thresholds. The outlines are overlaid on high-resolution Google satellite imagery (a and b) and a 2005 aerial image (c and d). Images b and d zoom in on the area of discussion, showing a ground moraine feature matching the RGI 7.0 terminus (b1), as well as a lake (d1), melt river (d2), possible ice cliffs (d3), and vegetation (d4).	64
Figure 27: Classification result of Belvedere glacier 2003 (Approach I) compared to RGI 7.0 outlines, overlaid on Landsat scene used for classification (a), high-resolution Google satellite imagery (b, c), and normalised LST (d).	66
Figure 28: Exemplary visualisation of the glacier tongue of Zmuttgletscher 2016 with (a) high-resolution Google satellite imagery, (b) Landsat 8 RGB of 30 m resolution, and (c) Landsat 8 LST [°K] of 100m resolution resampled to 30m, illustrating the challenges in DCG delineation with data of coarse resolution.....	68
Figure 29: Comparison of the Glacier Inventory outlines for Zmuttgletscher, showing whole glacier outlines for the years 2003 (RGI 7.0), 2010 (SGI2010) and 2016 (SGI2016). A narrowing of the glacier tongue can be observed for the outline of 2010.	69
Figure B 1: Classification results of Temporal changes, Unteraargletscher. The colorbar in the final classification sum shows the number of times a pixel was classified as DCG by the different band combinations.....	84
Figure B 2: Classification results of Temporal changes, Oberaletschgletscher.	84
Figure B 3: Classification results of Cross-glacier analysis, Approach I, Unteraargletscher.	85
Figure B 4: Classification results of Cross-glacier analysis, Approach I, Oberaletschgletscher.	85
Figure B 5: Classification results of Cross-glacier analysis, Approach I, Belvedere glacier.....	86
Figure B 6: Classification results of Cross-glacier analysis, Approach I, Satopanth glacier	86
Figure B 7: Classification results of Cross-glacier analysis, Approach II, Unteraargletscher.....	86
Figure B 8: Classification results of Cross-glacier analysis, Approach II, Oberaletschgletscher.....	87
Figure B 9: Classification results of Cross-glacier analysis, Approach III, Unteraargletscher.....	87
Figure B 10: Classification results of Cross-glacier analysis, Approach III, Oberaletschgletscher.....	88
Figure B 11: Classification results of Cross-glacier analysis, Approach III, Belvedere glacier.	88
Figure B 12: Classification results of Cross-glacier analysis, Approach III, Satopanth glacier.	88

Figure C 1: Overview of the input layers for the Random Forest Classification, for Zmuttgletscher 2016. All input layers were scaled to a value range of [0, 1]. The inclusion of Glacier outlines (current) as displayed here corresponds to the input dataset of temporal change analysis and cross-glacier analysis Approach I..... 90

Figure D 1: Creation of ground reference exemplary for Belvedere glacier 2003, with (a) mask of RGI 7.0 glacier outlines (with 0 = no glacier and 1 = glacier), (b) mask of the Normalised Difference Snow Index (NDSI) after applying a threshold of 0.4 (with 0 = no snow/ice and 1 = snow/ice), and (c) resulting ground reference after subtracting the NDSI mask from the RGI outlines (with 0 = no DCG and 1 = DCG)..... 90

Figure E 1: Complete comparison of the bands tested in a normalised difference index (NDI) with LST, showing the grayscales of each band, the normalised difference index with LST, and a scatterplot of the relation between the respective band (y-axis) and LST (x-axis), with data points coloured as non-DCG (grey) and DCG (blue), according to the SGI2016 debris cover outlines..... 91

List of Tables

Table 1: Overview of the RGI 7.0 attribute table (RGI 7.0 Consortium, 2023) for the glaciers of interest, with acquisition date of the satellite imagery that served as a basis for DCG classification. COPDEM refers to the Copernicus DEM..... 24

Table 2: Overview of the attributes in the SGI2010 (Fischer et al., 2014) for the glaciers of interest..... 25

Table 3: Overview of the attributes in the SGI2016 (Linsbauer et al., 2021) for the glaciers of interest..... 25

Table 4: Overview of the Landsat scenes selected for analysis of DCG in 2003, 2010 and 2016. 27

Table 5: Overview of the input layers for the Random Forest classification. 31

Table 6: Overview of statistical metrics to evaluate the performance of the classifier for temporal changes, applied to Zmuttgletscher, Unteraargletscher and Oberaletschgletscher..... 41

Table 7: Overview of statistical metrics to evaluate the performance of the classifier for cross-glacier analysis, Approach I, applied to Zmuttgletscher, Unteraargletscher and Oberaletschgletscher in 2003, 2010 and 2016, and to Belvedere and Satopanth glaciers in 2003. 44

Table 8: Overview of statistical metrics to evaluate the performance of the classifier for cross-glacier analysis, Approach II, applied to Zmuttgletscher, Unteraargletscher and Oberaletschgletscher in 2010 and 2016..... 48

Table 9: Overview of statistical metrics to evaluate the performance of the classifier for cross-glacier analysis, Approach III, applied to Zmuttgletscher, Unteraargletscher and Oberaletschgletscher in 2003, 2010 and 2016, and to Belvedere and Satopanth glaciers in 2003. 50

Table 10: Cross-comparison of validation accuracies of Normalised Difference Indices with LST and one additional band, and normalised LST only, with simplified RF classification. 53

Table 11: Comparison of statistical metrics for Zmuttgletscher 2003 with cross-glacier analysis, Approach I, when using different thresholds, either based on Otsu's method (9) or manually selected (17). 63

1 Introduction

1.1 Motivation and state of research

Accurate and comprehensive glacier inventories are essential for understanding the impacts of climate change, managing water resources and assessing natural hazards (Biddle, 2015; Nunchhani et al., 2024). In response to climate change, recent glacier retreat has led to an increase in debris-cover on glaciers (Bolch et al., 2007).

Precise mapping of debris-covered glaciers (DCGs) is particularly important for monitoring and modelling glacier dynamics, as DCGs often behave differently from clean-ice glaciers in terms of melt processes and mass balance (Herreid & Pellicciotti, 2020). Accurate delineation of DCGs is essential for improving climate models, particularly in regions where glaciers serve as primary freshwater sources (Nunchhani et al., 2024), such as the Himalayas, Andes and Alps. These regions are heavily dependent on glacier-fed water systems for agriculture, drinking water and hydropower (Agrawala et al., 2003), making accurate mapping of DCGs critical for predicting future water availability.

In addition, DCGs contribute to sea-level rise through ice melt, highlighting the importance of accurately assessing their extent and behaviour. Despite their importance, DCGs are poorly represented in global sea-level rise models (Raper & Braithwaite, 2006; Scherler et al., 2011b). Understanding their thermal and structural characteristics is also critical for incorporating DCG mapping into hydrological models, which are essential for predicting changes in water resources and managing natural hazards. For example, accurate delineation of DCGs helps predict glacial lake outburst floods (GLOFs), a serious hazard in glaciated regions (Agrawala et al., 2003). Refining DCG mapping will improve risk assessments and support the development of better mitigation strategies for such hazards (Biddle, 2015; Nunchhani et al., 2024).

However, the presence of debris-cover on the glacier poses a significant challenge to remote sensing-based glacier mapping (Karimi et al., 2012). Previous attempts to automate the mapping of DCGs using multispectral data have been limited by the spectral similarities in the optical spectrum between supraglacial debris and surrounding features such as bedrock, moraines, and fluvial deposition (Biddle, 2015; Karimi et al., 2012; Nunchhani et al., 2024).

Existing datasets, such as the Randolph Glacier Inventory version 6 (RGI 6.0), have shown notable inaccuracies in delineating the extent of DCG (Attaullah et al., 2023; RGI 7.0 Consortium, 2023). The recently published Randolph Glacier Inventory version 7.0 (RGI 7.0) does not include specific updates regarding DCGs, leaving uncertainties regarding their accurate representation (Aguayo et al., 2023; Maussion et al., 2023). Furthermore, the RGI does not provide a specific class for DCG. Instead, DCG areas are included within the overall glacier outline, without distinguishing between DCG and clean-ice glacier areas.

Similarly, while the Swiss Glacier Inventory 2016 (SGI2016) provides relatively accurate data for Swiss glaciers and includes a specific debris-cover class, the greatest inaccuracies are observed in the delineation of DCGs (Linsbauer et al., 2021). This is mainly due to the inherent challenges of mapping debris-covered areas. However, the SGI2016 achieves relatively high accuracy by relying on extensive manual work by glaciologists and high-resolution 3D imagery (Linsbauer et al., 2021). While this approach provides reliable results, it is highly labour-intensive and costly, making it impractical to scale to a global scale. These challenges highlight the need for more automated and scalable methods to improve the accuracy and reliability of DCG inventories worldwide.

In recent studies, thermal data has been used as a mean to differentiate between periglacial debris (PGD) and supraglacial debris (SGD) (Nunchhani et al., 2024; Sharda & Srivastava, 2024). The goal is to discern PGD from SGD by leveraging the temperature variance between them. It is hypothesized that glacial debris maintains a cooler temperature due to the presence of underlying ice, providing a basis for discrimination (Bolch et al., 2007).

While promising, the application of thermal data for systematic mapping is still in its early stages and faces significant challenges, including coarse spatial resolution, mixed pixel effects, and external influences such as shading, snow cover, and seasonal variations (Kaushik et al., 2022). Recent advances have pursued multispectral approaches, combining thermal data with optical satellite imagery, with promising results for partially automated DCG mapping (Karimi et al., 2012; Kaushik et al., 2022; Mitkari et al., 2022). However, these methods have so far been constrained by limited accuracy and scalability, indicating the need for further refinement to improve their accuracy and applicability across diverse glacier regions.

1.2 Research objectives and questions

The goal of this thesis is to develop and test a methodology for refining DCG outlines by exploiting the thermal differences between SGD and PGD. The methodology is applied and evaluated on glaciers in different mountain regions worldwide, providing a basis for its potential extension to time series analyses and larger regional scales. The thesis addresses the following research questions:

- RQ1 How can land surface temperature data be used effectively to distinguish between periglacial and supraglacial debris in glaciated regions?
- RQ2 What is the potential for applying the developed methodology to refine glacier outlines and assess glacier changes of debris-covered glaciers across different regions, and what are its limitations in terms of scalability and robustness?
- RQ3 How can the developed methodology complement existing glacier inventories, such as the Randolph Glacier Inventory (RGI) or the Swiss Glacier Inventory (SGI), to improve the accuracy of DCG delineation?

The outcome of the thesis is a glacier mapping methodology that could serve as an application for the upcoming satellite missions (TRISHNA, LSTM and SBG) which will provide thermal infrared (TIR) data at an unprecedented spatial (50 to 60-meter ground sampling distance) and temporal (1 to 3 days) scale and are scheduled for launch by the end of the decade (ESA, 2020, 2022; Lagouarde et al., 2018).

The thesis output contributes to the scientific community in two ways: firstly, by advancing the use of LST data in the cryosphere, which has so far been limitedly explored; and secondly, by offering a mapping tool for the glaciological community.

1.3 Research aims and structure

The aim of this study is to develop a methodology to distinguish between SGD and PGD, thereby identify DCGs by exploiting differences in the thermal signal. To achieve this, satellite imagery from Landsat 5 and Landsat 8, including land surface temperature (LST) data, was combined with topographic information from digital elevation models (DEMs). Calibration and validation of the methodology was carried out using existing glacier inventories, namely the RGI and the SGI.

The primary study site for methodology development was Zmuttgletscher, a DCG in the Swiss Alps where previous studies have been conducted (Mölg et al., 2019). Zmuttgletscher benefits from extensive data availability, making it an ideal site for calibration and validation. To increase the extendibility of the methodology, two additional DCGs in the Swiss Alps, Unteraargletscher and Oberaletschgletscher, were included. All Swiss glaciers are catalogued in the Swiss Glacier Inventory (SGI), whose latest update in 2016 provides highly accurate outlines and includes a separate classification for DCGs, which further supported the calibration and validation efforts.

The methodology was developed within the Google Earth Engine (GEE) environment, taking advantage of its ability to directly import satellite imagery. The process drew on insights from previous studies using TIR and LST data to map DCGs but was largely developed through iterative experimentation and testing of different configurations. Once the classification showed satisfactory performance in capturing temporal changes on Zmuttgletscher, Unteraargletscher and Oberaletschgletscher for three specific years (2003, 2010 and 2016), all of which had corresponding glacier inventory data, it was applied to analyse its performance across different glaciers.

After calibrating the methodology on the Swiss glaciers, it was extended to two different glaciers on which the model had not been trained: Belvedere in the Italian Alps and Satopanth in India. This approach allowed testing in regions with different climatic conditions, further assessing the adaptability and robustness of the methodology in different environmental contexts.

This thesis is structured as follows: first, a theoretical background section provides essential information on thermal satellite imagery and its application in DCG mapping. Next, the study areas are introduced, followed by a detailed data section describing the datasets used in this study. The methods section then outlines the development of a methodology for refining DCG outlines, addressing RQ1. The results section presents key findings from the application of this methodology to Zmuttgletscher, Unteraargletscher, Oberaletschgletscher, Belvedere glacier and Satopanth glacier, including mapped outlines and statistical performance metrics. This is followed by a discussion of the potential and implications of the developed methodology, addressing RQ2 and RQ4. A detailed discussion of its limitations will address RQ3. The thesis concludes with a summary of the main findings and an outlook on potential directions for further research in this area.

2 Theoretical Background

2.1 Thermal infrared imagery and advancing LST retrieval

Thermal infrared (TIR) remote sensing, operating within the 8–14 μm atmospheric window wavelength region, has been an essential tool in Earth observation since the 1960s (Li et al., 2023). TIR instruments detect the thermal radiation emitted by the Earth's surface, enabling the estimation of land surface temperature (LST), which reflects how hot or cold the Earth's surface would feel to the touch (Dash et al., 2002; Li et al., 2023). LST is influenced by factors such as vegetation, soil moisture, and surface materials, making it a critical parameter for understanding various environmental processes, including climate patterns (Bechtel, 2015), urban heat islands (Alexander, 2020), and glacier dynamics (Nunchhani et al., 2024).

Since the launch of early satellites like NASA's Nimbus in the 1960s, a wide array of satellites has carried TIR sensors, including the Landsat series, Earth Observing System (EOS) Terra and Aqua satellites, Meteosat Second Generation (MSG) satellite series, and the Chinese Fengyun (FY) satellite series (Li et al., 2023). These TIR sensors have been widely used to generate different LST products across various spatial resolutions, from regional to global scales (Malakar et al., 2018).

Over the past decade, there has been growing interest in LST retrieval from TIR data, leading to the development of new LST retrieval algorithms, the launch of new TIR instruments, and the release of updated LST products (Li et al., 2023). The process of retrieving LST from TIR data, however, presents several challenges. LST is determined by measuring radiances influenced by surface parameters and the atmospheric composition and thermal structure. This makes LST retrieval a complex problem, requiring various assumptions and constraints for accurate estimation (Li et al., 2023).

Several retrieval algorithms have been developed since the 1970s to address these challenges, aiming to retrieve accurate LST by accounting for atmospheric effects and surface emissivity. They use radiative transfer principles to correct for distortions caused by the atmosphere but differ in their complexity and data requirements: the Single-Channel (SC) algorithm is straightforward and uses one thermal band, making it computationally simple but less robust under varying atmospheric conditions. The Split-Window (SW) algorithm improves accuracy by leveraging the difference between two thermal channels, whereas the Temperature and Emissivity Separation (TES) algorithm goes a step further by estimating emissivity alongside temperature using spectral data. The Day/Night (D/N) algorithm uniquely exploits diurnal temperature variation for greater emissivity detail but requires data from both day and night passes, which can limit its applicability (Li et al., 2023; Malakar et al., 2018).

Despite these advancements, current LST products face several limitations. For instance, spatial discontinuity occurs due to the inability of TIR remote sensing to penetrate clouds, leading to missing LST values over cloud-covered pixels (Bechtel, 2015).

Generating LST requires accurate land surface emissivity, which is provided by NASA's Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Emissivity Dataset (GED). However, gaps in the ASTER GED dataset, caused by the exclusion of pixels with unrealistic emissivity values, result in missing LST data. As a result, certain areas lack LST values, limiting the completeness of thermal information for these regions (Hulley & Hook, 2015).

Additionally, there is a lack of spatiotemporal comparability, as differences in local viewing time and angle can result in significant variations in LST for the same pixel on different days. Other challenges include the short time spans covered by many LST products (Bechtel, 2015), which limits their use in

long-term analysis, and the instantaneity of LST measurements, which only provide data at the satellite overpass time (Li et al., 2023).

The upcoming thermal missions, such as TRISHNA (Thermal InfraRed Imaging Satellite for High-resolution Natural resource Assessment) (Lagouarde et al., 2018), LSTM (Land Surface Temperature Monitoring) (Koetz et al., 2018), and SBG (Surface Biology and Geology) (Schimel & Poulter, 2022), scheduled for launch by the end of the decade, are expected to address some of these limitations. These missions will offer unprecedented spatial resolutions of up to 60 meters and temporal resolutions of 3 days (Koetz et al., 2018; Lagouarde et al., 2018; Schimel & Poulter, 2022), significantly enhancing the ability to monitor and study thermal processes of the Earth, including the thermal characteristics of DCGs. These advancements will likely lead to more accurate and reliable LST products, contributing to a deeper understanding of environmental dynamics.

2.2 The use of thermal data in DCG mapping

Due to a cooling effect of the underlying ice, SGD is assumed to show a lower temperature than PGD (Bolch et al., 2007). By leveraging the temperature contrasts in TIR data arising from the presence or absence of underlying ice, the inclusion of thermal data addresses the challenge posed by the similarities of SGD and PGD in the optical spectrum (Figure 1). While various studies have approached this issue differently, recent trends favour deep learning models for automating DCG mapping (Karimi et al., 2012; Kaushik et al., 2022). However, these approaches face challenges related to data resolution and factors beyond ice presence, such as terrain aspect (shaded or illuminated areas), which influence LST. Moreover, threshold-based methods can limit applicability across diverse regions, affecting model generalisation.

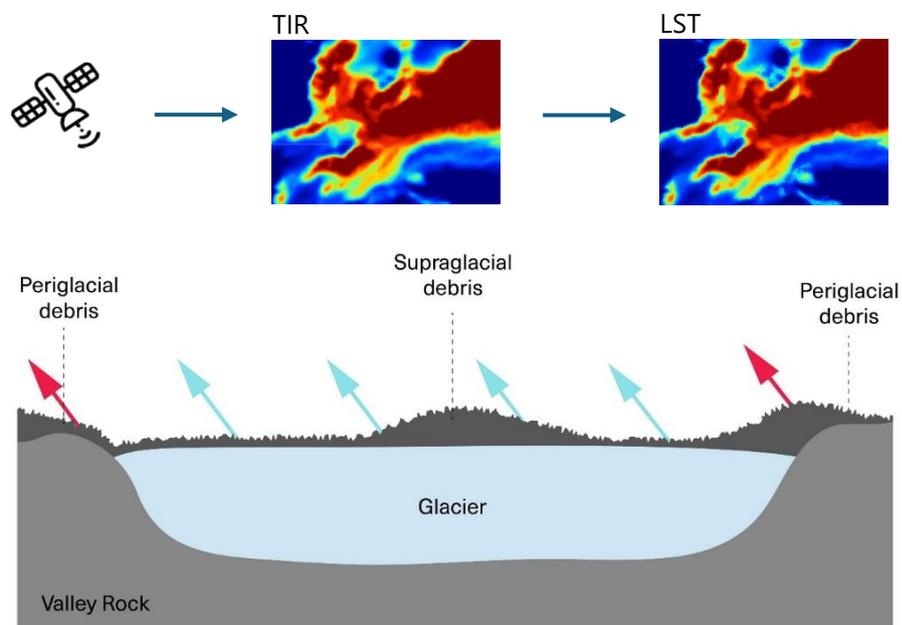


Figure 1: Illustration of a schematic cross-section of a DCG based on Shukla et al. (2010), and the retrieval of LST to distinguish between PGD and SGD.

Several studies have explored the potential of thermal data in mapping DCGs, often combined with additional datasets to improve accuracy and overcome the challenges posed by complex terrain and

spectral similarity between SGD and surrounding PGD. TIR data, particularly from Landsat's Thematic Mapper (TM) bands, have been a major focus, as shown by Karimi et al. (2012), who used temperature differences between SGD and PGD for delineation. They achieved high accuracy using a multi-source approach, including optical, thermal and laser scanning (LiDAR)-derived DEM data. However, limitations in spatial resolution and the non-open-source nature of the WorldView-2 imagery used highlight barriers of the extensibility of this approach.

Recent advances have focused on automation and machine learning. Kaushik et al. (2022) developed a deep neural network that integrated a wide range of datasets, including optical, near-infrared (NIR), short-wave infrared (SWIR), TIR, microwave, and elevation data, to produce robust results over the Himalayas and Karakoram regions. Despite the promise of such automated methods, challenges remain, such as misclassification of shadows as debris and confusion between dirty ice, snow and SGD. These errors highlight the difficulty of identifying DCGs in heavily debris-laden areas, even with manual corrections. Kaushik et al. (2022) suggested incorporating additional data sets, such as surface velocity data, to address these limitations.

Other studies have combined rule-based methods with object-based image analysis (OBIA) to refine DCG delineation. For example, Mitkari et al. (2022) utilised thermal, optical and DEM-derived slope data with a focus on differentiating SGD from PGD by incorporating slope thresholds and segmentation techniques. While this Approach is promising, it requires extensive manual adaptation of thresholds for each region, which limits scalability. They recommended machine learning to improve automation and regional applicability.

Across these studies, the trend is towards integrating diverse datasets and using automated approaches to improve the accuracy and scalability of DCG mapping. However, challenges remain, particularly related to the limited spatial resolution of thermal data (Karimi et al., 2012; Kaushik et al., 2022; Mitkari et al., 2022), the need for region-specific adjustments (Mitkari et al., 2022), and the reliance on non-open source datasets such as WorldView-2 imagery (Karimi et al., 2012). These limitations highlight the need for methods that use freely available data and are adaptable to different regions, ensuring wider applicability and filling gaps in current global glacier inventories.

Whilst building on the findings of these previous studies, this work aims to maximise the extensibility of the methodology across different regions of interest, avoiding threshold dependency and relying solely on open-source data.

2.3 Influence of debris thickness on LST

Debris thickness on glaciers plays a critical role in influencing LST and the detection of underlying ice. On DCGs, debris generally moves downstream, with increasing debris thickness towards the glacier terminus (Mölg et al., 2020; Nakawo et al., 1986). When the debris-cover is thinner than 30-40 cm, the lower temperatures of the underlying ice make it detectable in thermal satellite imagery (Ranzi et al., 2004). However, as debris thickness increases, this cooling effect diminishes, making thermal detection more difficult. Studies show that debris thicker than 40-50 cm no longer reflects the cooling influence of the underlying ice, making it difficult to identify using thermal data (Biddle, 2015; Ranzi et al., 2004; Taschner & Ranzi, 2002).

Debris thickness data are not available for most DCG, as field measurements are very labour intensive due to the large variability even within a single glacier. For example, measurements on Zmuttgletscher

in the Swiss Alps show debris thicknesses ranging from less than 5 cm to more than 70 cm, illustrating significant variability (Mölg et al., 2019). This highlights the need for robust remote sensing techniques, as debris thickness significantly affects the thermal signal in satellite imagery.

2.4 Additional bands used in DCG mapping

Thermal satellite data (TIR or LST) alone has proven insufficient for effective DCG mapping due to limitations in spatial resolution and sensitivity to external factors. The coarse resolution of thermal data often results in mixed pixels, where multiple surface types, such as debris, vegetation and clean ice, are captured within a single pixel, reducing the reliability of the thermal signal. In addition, thermal data are influenced by topographic effects such as slope and aspect, and environmental factors such as shading, snow cover or surface moisture, further complicating their usefulness in discriminating between SGD and PGD (Alifu et al., 2015; Jawak et al., 2022; Kaushik et al., 2022). These challenges highlight the need to integrate LST with other datasets to improve accuracy and address its inherent limitations.

Studies have shown the effectiveness of combining TIR with different spectral bands to improve DCG mapping (Jawak et al., 2022; Kaushik et al., 2022). For example, visible and near-infrared (VNIR) bands are particularly useful for identifying dirty ice, as their reflectance properties vary significantly between different surface types (Jawak et al., 2022). Healthy vegetation reflects strongly in the NIR, while rock and debris have low to moderate reflectance, depending on factors such as composition and moisture content (Holzman et al., 2021). On glaciers, clean ice and snow also reflect NIR effectively, but this reflectance decreases with contamination by debris or melting processes (Pope & Rees, 2014). In addition, shortwave infrared (SWIR) bands have been used to analyse the mineral composition of debris, complementing VNIR data for detailed mapping of glacier facies (Jawak et al., 2022).

Thermal bands have been shown to further enhance debris detection when combined with reflective properties in the NIR and SWIR. Alifu, Tateishi, and Johnson (2015) developed a band ratio combining TIR, NIR, and SWIR bands for DCG mapping, demonstrating how the integration of thermal and reflective data can improve the distinction between SGD and PGD. The utility of NIR and SWIR bands lies in their ability to capture the distinct spectral characteristics of glacier surfaces and surrounding features, providing additional layers of information to complement LST-based analyses.

These advances highlight the value of multispectral approaches that combine thermal and reflective data to overcome the limitations of LST alone. By using complementary data sets, DCG delineation can be refined, and the reliability of mapping methods is improved.

3 Study Sites

The methodology was developed with Zmuttgletscher as the primary study area, and calibration was performed using data from Zmuttgletscher, Unteraargletscher and Oberaletschgletscher in the Swiss Alps to ensure a robust dataset. The analysis was then extended to Belvedere Glacier in northern Italy and Satopanth Glacier in the Himalayas, northern India, to assess the performance of the method in different climates and data availability.

The application of the methodology to other Swiss glaciers benefited from the use of Swiss Glacier Inventory (SGI) datasets, which were also used for the calibration and validation of Zmuttgletscher. This allowed direct comparison of model performance across Swiss glaciers under similar climatic conditions, glacier sizes, elevation ranges and consistent standards for ground reference data.

The extension to Belvedere Glacier in the Italian Alps provided a glacier with comparable climatic conditions but was limited to the 2003 Randolph Glacier Inventory (RGI) for ground reference data. While the availability and accuracy of the ground reference data differed from the Swiss glaciers used in the initial analysis, the consistency of the remaining data set and framework allowed an assessment of the model's applicability under slightly different data conditions in comparable environmental settings.

Finally, the methodology was extended to Satopanth Glacier in the Himalayas, again using RGI data for comparison. This application tested the model under substantially different climatic conditions, allowing an assessment of its robustness and adaptability beyond the European Alps.

3.1 Zmuttgletscher

The primary study site was Zmuttgletscher [45°59' N, 7°37' E] in the Matter Valley of the western Swiss Alps (Figure 2). The extensive data availability and accessibility make Zmuttgletscher an ideal study area for calibration and validation of the method.

Spanning an elevation from approximately 2240 to 4150 meters above sea level (m asl), the glacier is flanked by the Matterhorn (4478m) and the Dent d'Hérens (4174m) to the south, and the Dent Blanche (4357m) to the north, which supply the glacier system with debris and avalanches. In 2016, according to Mölg et al (2019), the glacier spanned an area of 15.74 km², with a substantial portion covered by debris originating from the surrounding rock walls. The SGI assumes a slightly smaller area of 15.82 km² in the same year (GLAMOS, 2023). Zmuttgletscher lies in a relatively dry region at the main divide of the Alps, receiving precipitation from both northern and southern weather systems. While there are no direct measurements at higher elevations, estimates suggest accumulation values between 0.8 and 1.5 meters (Mölg et al., 2019). The glacier has several tributaries, including Tiefmattengletscher to the south, Stockjigletscher to the west, and Schönbielgletscher to the north. Historically, the main glacier tongue was nourished by all these accumulation areas, but in recent years, Tiefmattengletscher and, to a lesser extent, Stockjigletscher have been the main contributors. Around 2010, the almost debris-free central Stockjigletscher branch detached from the main tongue, while Schönbielgletscher maintains a continuous debris-cover even above the icefall at approximately 2900 meters (Mölg et al., 2019). Field measurements on Zmuttgletscher indicate debris thicknesses ranging from less than 5 cm to over 70 cm, with thicker deposits typically found on elongated ridges and steeper slopes (Mölg et al., 2019).

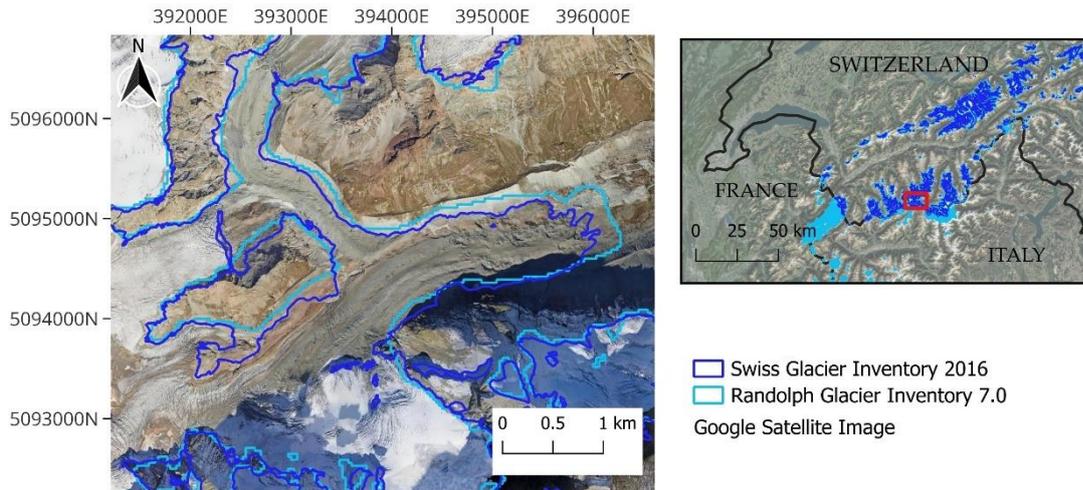


Figure 2: Overview of Zmuttgletscher as the primary study site and its location in the Swiss Alps, showing high-resolution Google satellite imagery overlaid with the SGI2016 glacier outlines.

3.2 Unteraargletscher

Unteraargletscher is the fourth-largest glacier in the Swiss Alps in terms of both area and length. In 2016, the glacier covered an area of 22.7 km² and stretched over 12.3 km (GLAMOS, 2022). Located in the Bernese Alps, west of the Grimsel Pass, the glacier resembles a Y-shape, formed by the confluence of two glacier arms (Figure 3). The southeasterly tributary, Lauteraargletscher, originates at an elevation of 4017 m asl between Schreckhorn and Bärglistock. The northeasterly tributary, Finsteraargletscher, begins at the flanks of Agassizhorn, Finsteraarhorn, and Oberaarhorn, with the Strahlegg Glacier also flowing into the Finsteraargletscher. The merging of these glaciers creates a large medial moraine. Both glacier arms are roughly 1 km wide and have a similar surface inclination of approximately 4–5°. The glacier is flanked by steep side walls (Bauder et al., 2003; GLAMOS, 2022).

The terminus of Unteraargletscher has a relatively low slope of about 4° and is almost entirely covered by debris. The debris layer is typically 10 to 20 cm thick, with the thickness increasing as it approaches the glacier terminus. By 2016, the glacier tongue had retreated to an elevation of 1931 m, situated 1.5 km from Grimselsee (GLAMOS, 2022; Huss et al., 2007).

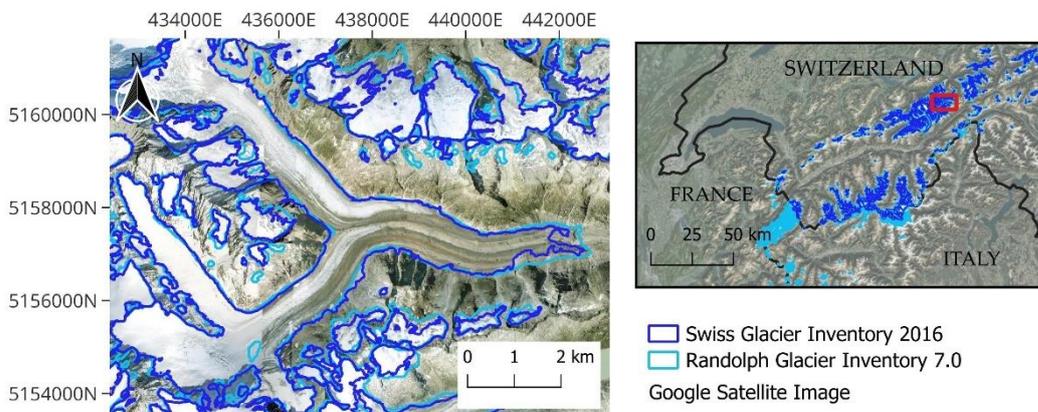


Figure 3: Overview of the study site of Unteraargletscher and its location in the Swiss Alps.

3.3 Oberaletschgletscher

Oberaletschgletscher is located on the southern side of the Bernese Alps in the canton of Valais (Figure 4). In 2011, the glacier covered an area of 17.47 km² with a length of 9.16 km (GLAMOS, 2022). In comparison, the glacier's area in 1973 was 21.62 km², indicating a 19.17% reduction in surface area over four decades, reflecting a clear trend of retreat consistent with other Alpine glaciers (GLAMOS, 2022).

Oberaletschgletscher is a valley glacier that flows from north to southeast (Langhammer et al., 2019). Its terminus is at an elevation of 2130 m asl, making the glacier particularly sensitive to climatic changes (Peña-Haro et al., 2021). The glacier tongue is heavily debris-covered (Jouvet et al., 2011), with supraglacial material originating mainly from the northern headwalls and a tributary glacier to the west (Paul et al., 2004). This extensive debris-cover plays a crucial role in influencing glacier dynamics and response to climate change (Langhammer et al., 2019; Mölg et al., 2020).

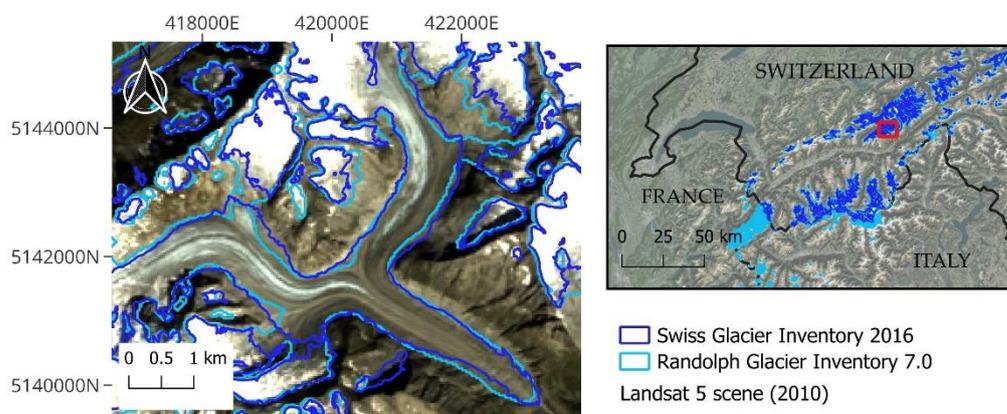


Figure 4: Overview of the study site of Oberaletschgletscher and its location in the Swiss Alps. Since the high-resolution Google Satellite Image of this area was snow covered, a scene of Landsat 5 (2010) was used for visualisation.

3.4 Belvedere Glacier

Belvedere glacier is a DCG located northeast of the highest peaks of the Monte Rosa massif, in the northwestern Italian Alps (Figure 5). Due to its debris-cover and favourable solar exposure, the glacier's frontal sectors extend to relatively low elevations, terminating at an altitude of 1785 m asl in the year 2000 (Ranzi et al., 2004). The glacier is the terminus of four higher glaciers: Nordend, Monte Rosa, Signal and Northern Locce glaciers (Colombero et al., 2019). In the RGI 7.0, Belvedere glacier is outlined as both the more elevated glacier areas and the lower, debris-covered part, resulting in a total area of 4.45 km² in 2003 (RGI 7.0 Consortium, 2023).

Measurements carried out by the Italian Glaciological Committee (CGI) in 2006 showed that the debris-covered area of Belvedere glacier had a surface area of 1.46 km² and a maximum length of 3091 metres (CGI-CNR (Comitato Glaciologico Italiano & Consiglio Nazionale delle Ricerche), 2024). In 2006, the glacier covered an altitude range from 2397 to 1770 m asl with an average slope of 8°. The terminus of the glacier has a bilobate structure, with both lobes showing signs of retreat (Salvatore et al., 2015).

In 2019, the larger northern lobe had an average length of 650 metres and reached a minimum elevation of about 1810 metres, 40 metres higher than in 2006. The southern lobe was 350 metres long and terminated at an elevation of 1840 metres above sea level. The two lobes are separated by a medial moraine (Colombero et al., 2019).

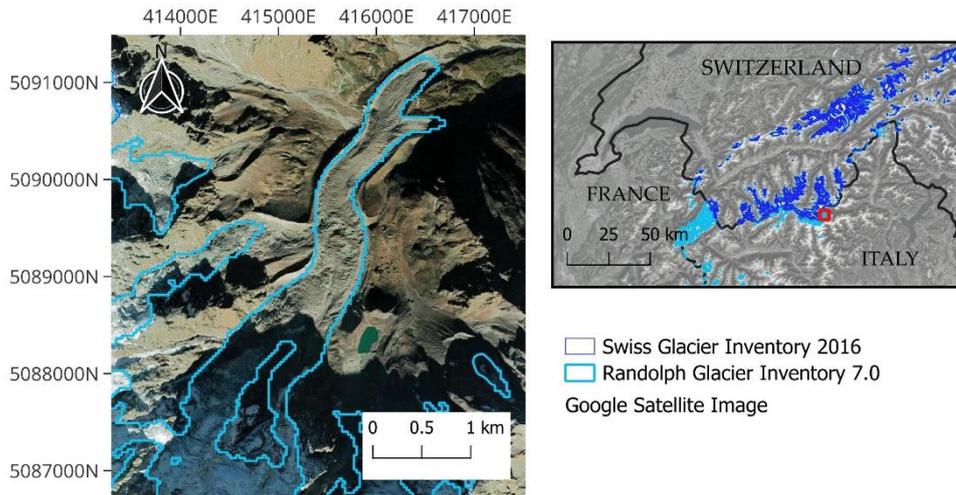


Figure 5: Overview of the study site of Belvedere glacier and its location in the Italian Alps.

3.5 Satopanth Glacier

Satopanth glacier, located in northern India (Figure 6), covers an area of approximately 19 km², with approximately 60% of its surface area being covered by debris (Shah et al., 2019). The glacier extends over an elevation range of 3900 to 6200 m, with the debris-cover starting at around 4500-4700 m and spanning roughly 800 m in elevation. The debris-cover, derived mainly from the weathering of steep headwalls and sidewalls, is up to one metre thick and covers an area of about 11 km² (Shah et al., 2019).

In the debris-covered section, the glacier slope is generally gentle, while the clean ice area above 4700 m remains steep (Nainwal et al., 2016). During 2015-2017, subdebris ablation rates on Satopanth glacier ranged from 1.5 to 1.7 cm per day (Shah et al., 2019).

Records show that the glacier has been retreating since at least 1936, with an average frontal retreat of ~6 m per year and a relatively stagnant lower ablation zone with ice flow rates below 5 m per year. This zone has thinned at a rate of ~0.4 m per year over the past 50 years (Nainwal et al., 2016). These are typical characteristics of DCGs in the Himalayas, making Satopanth representative of its type (Scherler et al., 2011a). The application of the methodology in the Himalayas provides an opportunity to assess its performance in different climatic conditions and regional contexts.

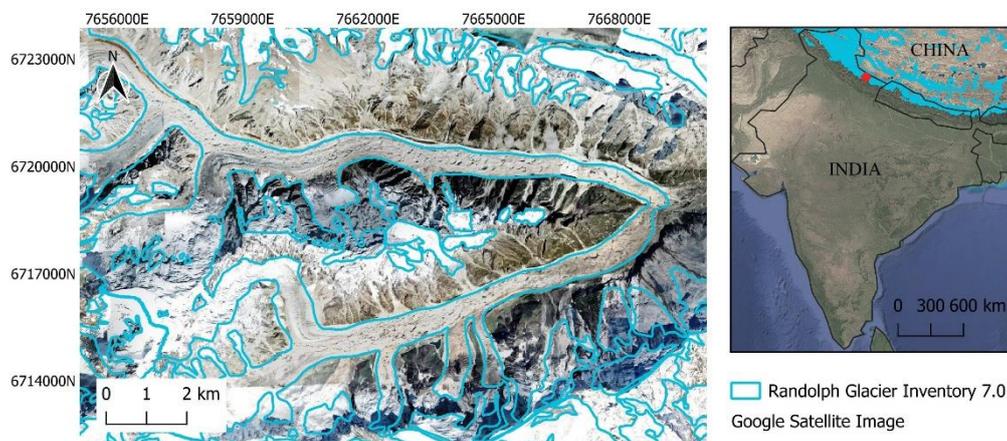


Figure 6: Overview of the study site of Satopanth Glacier and its location in northern India, Himalaya.

4 Data

This thesis used LST data to develop and validate a methodology for refining the outlines of DCGs. To achieve this, the analysis integrated three primary data sources: satellite imagery, Digital Elevation Models (DEMs), and established glacier inventories.

High-resolution satellite imagery from Landsat 5 and Landsat 8 provided LST data and additional spectral information at other wavelengths. Complementary topographic data from DEMs derived from NASADEM, Copernicus and SwissAlti3D captured terrain details over several years. In addition, three glacier inventories, the Randolph Glacier Inventory version 7 (RGI 7.0) and the Swiss Glacier Inventories (SGI) of 2010 and 2016 provided essential baselines for validation and calibration.

Satellite imagery and DEM-derived topographic information supported the development of the method, while the glacier inventories were essential for calibration and validation. Together, this multi-source dataset provided a robust basis for accurate refinement of glacier contours, increasing the reliability of results across multiple glaciers and years.

4.1 Satellite Imagery

Landsat data was utilized to conduct this study, specifically drawing from Landsat 8 for the year 2016 and Landsat 5 for the years 2003 and 2010. The collection 2, Tier 1 Level 2 (C02/T1_L2) dataset was employed, which contains atmospherically corrected surface reflectance and LST.

The optical and near- to shortwave-infrared bands have a spatial resolution of 30 meters (Li et al., 2023). Originally at 100 meters resolution, LST is resampled to 30 meters to match the spatial resolution of the other bands in this product (Li et al., 2023; Malakar et al., 2018).

The temporal resolution of the Landsat data is 16 days (Li et al., 2023). However, in mountainous regions, the effective use of this temporal resolution is often compromised by shadows, clouds, and seasonal snow cover, significantly reducing the amount of usable imagery for analysis.

All Collection 2 surface temperature products are created with a single-channel (SC) algorithm jointly developed by the Rochester Institute of Technology (RIT) and the National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL) (Barsi et al., 2014; Jimenez-Munoz et al., 2009). The collected data are packaged into overlapping scenes that cover approximately 170 km x 183 km, using a standardized reference grid (USGS, 2021). For successful processing to surface temperature, data products must contain both optical and thermal data, as ASTER NDVI is required to temporally adjust the ASTER Global Emissivity Database (GED) product to the target Landsat scene (Hulley & Hook, 2015; Malakar et al., 2018); consequently, nighttime acquisitions cannot be processed for surface temperature (USGS, 2021).

4.1.1 Landsat 5

For the years 2003 and 2010, the study utilized the USGS Landsat 5, Collection 2, Tier 1 Level 2 dataset (LANDSAT/LT05/C02/T1_L2, Landsat 5 Imagery courtesy of the U.S. Geological Survey). This dataset includes atmospherically corrected surface reflectance and land surface temperature derived from the Landsat Thematic Mapper (TM) I sensor (Markham et al., 2004). It contains four visible and near-infrared (VNIR) bands and two shortwave infrared (SWIR) bands, all processed to orthorectified surface

reflectance, as well as one TIR band processed to orthorectified surface temperature. The dataset also includes intermediate bands utilized in the calculation of surface temperature products, alongside quality assurance (QA) bands (Acharya & Yang, 2015).

Landsat 5 surface reflectance products are generated using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm (version 3.4.0) (Schmidt et al., 2013).

4.1.2 Landsat 8

For the analysis of years later than 2013, this study utilised imagery of Landsat 8, specifically from the collection LANDSAT/LC08/C02/T1_L2 (Landsat 8 Imagery courtesy of the U.S. Geological Survey).

Landsat 8, equipped with the Operational Land Imager (OLI) and the TIR Sensor (TIRS), acquires data in the optical and thermal spectrum (USGS, 2021). In particular, bands 10 and 11, which measure radiation in the intervals 10.6-11.19 μm and 11.50-12.51 μm , allow for the retrieval of LST. Thermal data is available at a resampled resolution of 30m, offering seasonal coverage of the global landmass since February 2013 (Landsat 8) (Barsi et al., 2014; Earth Resources Observation and Science (EROS) Center, 2020).

The Landsat 8 dataset includes atmospherically corrected surface reflectance and LST, encompassing five visible and near-infrared (VNIR) bands and two shortwave infrared (SWIR) bands, all processed to orthorectified surface reflectance, along with one TIR band processed to orthorectified surface temperature. Additionally, like Landsat 5, it contains intermediate bands used in the calculation of surface temperature products and quality assurance (QA) bands (Acharya & Yang, 2015).

Landsat 8 surface reflectance products are produced using the Land Surface Reflectance Code (LaSRC) (Skakun et al., 2019). LST is generated from several input data sources, including atmospheric profiles and the NASA Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Emissivity Dataset (GED) (Hulley & Hook, 2015).

4.2 Digital Elevation Models (DEM)

Three different Digital Elevation Models (DEMs) were used in this study to account for topographic variation and to normalise LST values. The DEMs were chosen to match the analysis years as closely as possible, recognising that changes in glacier extent over time alter the topography. In the absence of annual DEM updates, the closest available DEMs to the years of analysis were selected to minimise discrepancies and ensure accurate normalisation of the LST data.

4.2.1 SwissAlti3D

SwissAlti3D is a high-resolution DEM provided by the Federal Office of Topography (Swisstopo). It offers topographic data at 0.5- and 2-meter resolution, providing highly detailed information on elevation changes, including the impact of glacial retreat and other landscape alterations. SwissAlti3D is produced from airborne laser scanning (LiDAR) and photogrammetric methods, ensuring high accuracy in steep and variable terrain. The data is processed and orthorectified to remove geometric distortions, resulting in a georeferenced dataset that is suitable for various analytical purposes, including glacier and topography studies (Bundesamt für Landestopografie swisstopo, 2022). The version that was used in this

study stems from the year 2019 and was used for the analyses of glaciers in Switzerland for the year 2016, at a resolution of 2 meters.

4.2.1 NASADEM

The NASADEM is a digital elevation model originating from the Shuttle Radar Topography Mission (SRTM), which provides elevation data with a resolution of 30 meters (NASA JPL, 2020). The SRTM data was collected using interferometric radar technology during a single 11-day mission in 2000, aboard the Space Shuttle Endeavour. The resulting elevation data was processed to fill gaps and remove artifacts, making it suitable for global applications (Buckley et al., 2020). The NASADEM was used in this study for the analysis of the year 2003.

4.2.2 Copernicus DEM

The Copernicus DEM provides global coverage with a resolution of 30 meters (Copernicus, 2010-2018). This DEM is part of the Copernicus Earth Observation Programme, funded by the European Union, and is based on data from the TerraSAR-X and TanDEM-X satellite missions, managed by the German Aerospace Center (DLR). The DEM is produced using synthetic aperture radar (SAR) interferometry, which allows for accurate elevation measurements by detecting the phase difference between radar signals received from slightly different angles. The data undergoes further processing to ensure consistency and global coverage, including noise filtering and the elimination of any anomalies introduced during data acquisition. This makes the Copernicus DEM particularly useful for large-scale applications, such as global terrain analysis and environmental monitoring (Fahrland et al., 2022).

The acquisition date of the Copernicus DEM that was used in this study stems from the period of 2011-2015 and was used for analysis of the year 2010.

4.2.3 SwissAlti3D

SwissAlti3D is a high-resolution DEM provided by the Federal Office of Topography (Swisstopo). It offers topographic data at 0.5- and 2-meter resolution, providing highly detailed information on elevation changes, including the impact of glacial retreat and other landscape alterations. SwissAlti3D is produced from airborne laser scanning (LiDAR) and photogrammetric methods, ensuring high accuracy in steep and variable terrain. The data is processed and orthorectified to remove geometric distortions, resulting in a georeferenced dataset that is suitable for various analytical purposes, including glacier and topography studies (Bundesamt für Landestopografie swisstopo, 2022). The version that was used in this study stems from the year 2019 and was used for the analyses of glaciers in Switzerland for the year 2016, at a resolution of 2 meters.

4.3 Glacier Inventories

To assess the performance of the developed methodology, its results were compared to existing glacier inventories. The Randolph Glacier Inventory (RGI) provides global coverage of glacier outlines, while the Swiss Glacier Inventories (SGI) served as an additional data source specifically for glaciers in the Swiss Alps.

4.3.2 Randolph Glacier Inventory version 7.0 (RGI 7.0)

The Randolph Glacier Inventory version 7.0 (RGI 7.0) is an extensive dataset representing glacier outlines and attributes on a global scale. It provides a thorough representation of global glacier distribution, excluding ice sheets in Greenland and Antarctica, and is compiled from various remote sensing data sources (RGI 7.0 Consortium, 2023).

The RGI is a subset of the Global Land Ice Measurements from Space (GLIMS) database and provides a snapshot of global glacier distribution around the year 2000. The RGI includes all glaciers larger than 0.01 km² and focuses on global coverage and consistency, rather than precise outline delineation and measurement of glacier area change. Its strength lies in enabling large-scale analyses, such as glacier volume estimations and regional response of the cryosphere to climate change (RGI 7.0 Consortium, 2023).

Developed through an international effort since 2010, the RGI is managed by the IACS Working Group on the RGI and Infrastructure for Glacier Monitoring. The inventory is hosted on the GLIMS platform, with data freely available (RGI 7.0 Consortium, 2023). The RGI 7.0 is an improved form of the RGI version 6.0 (RGI 6.0) and is based Landsat 5 TM images and Sentinel-2 images, and very high-resolution images from the "World imagery" layer of the ESRI Basemap (RGI 7.0 Consortium, 2023). In this study, the RGI 7.0 was used to assess the methods performance in 2003 and its applicability on a global scale. An overview of the RGI attributes for the glaciers of interest is provided in Table 1.

Table 1: Overview of the RGI 7.0 attribute table (RGI 7.0 Consortium, 2023) for the glaciers of interest, with acquisition date of the satellite imagery that served as a basis for DCG classification. COPDEM refers to the Copernicus DEM.

Glacier name	Acquisition date	UTM zone	Area [km ²]	Elevation median [m]	Slope [°]	Aspect [°]	DEM source	Maximum length [m]
Zmutt	13.08.2003	32	15.4	2984.1	18.8	36.4	COPDEM90	7784
Unteraar	13.08.2003	32	23.6	2648.2	18.1	65.9	COPDEM90	12530
Oberaletsch	13.08.2003	32	19.3	2964.7	21.8	122.1	COPDEM90	9659
Belvedere	13.08.2003	32	4.5	2835.2	28.6	42.5	COPDEM30	6326
Satopanth	10.07.2002	44	57.0	4901.0	19.7	34.4	COPDEM90	20106

4.3.2 Swiss Glacier Inventory 2010 (SGI2010)

The Swiss Glacier Inventory 2010 (SGI2010) is based on high-resolution (0.25 - 0.50 m) aerial orthophotos (SWISSIMAGE) acquired between 2008 and 2011. The inventory was manually delineated by a single expert, using the high-resolution orthophotos and the SGI1973 as a reference data set. To assess the accuracy of the approach, the manually delineated outlines were compared to an independent experiment on multiple digitization of clean, snow- and/or DCGs. Showing overall very robust results, the accuracy decreases for small glaciers (>1.0 km²) and for DCG, systematically underestimating these glacier areas (Fischer et al., 2014).

The SGI2010 contains outlines of all glaciers in the Swiss Alps (Fischer et al., 2014; Linsbauer et al., 2021), however without distinguishing between DCG and non DCG. Table 2 gives an overview of the attributes for the glaciers of interest.

Table 2: Overview of the attributes in the SGI2010 (Fischer et al., 2014) for the glaciers of interest.

Glacier name	sgi-id	Area [km ²]
Zmutt	B57-05	13.7
Unteraar	A54g-11	22.5
Oberaletsch	B36-01	17.5

4.3.1 Swiss Glacier Inventory 2016 (SGI2016)

The SGI2016 dataset provides detailed glacier outlines, including supraglacial debris-cover and ice divides, for all Swiss glaciers in a state between 2013 and 2018. It represents the most accurate and comprehensive cartographic representation of glacier extent in Switzerland to date (Linsbauer et al., 2021).

The Swiss Glacier Inventory 2016 (SGI2016) dataset was produced by the Federal Office of Topography and is regularly maintained and updated. It is based on various input data such as aerial orthophoto mosaics (SWISSIMAGE), the swissAlti3D, the Topographic Landscape Model (swissTLM3D) and its object classes 'debris' and 'glacier'. Both the swissTLM3D and the SGI2016 are based on manual outline delineation by glaciologists. In an expert workshop to define the requirements for the new SGI2016, a guideline was developed to enable the development of a consistent and exact glacier inventory (Linsbauer et al., 2021). The SGI2016 includes detailed mappings of glaciers within Switzerland, as well as separate mappings of DCGs, which served as ground reference for evaluating the performance of the new methodology for the year 2016 for the three Swiss glaciers (Table 3).

Table 3: Overview of the attributes in the SGI2016 (Linsbauer et al., 2021) for the glaciers of interest.

Glacier name	sgi-id	Year of acquisition	Area [km ²]	Length [km]	Elevation median [m]	Slope [°]	Aspect [°]	Debris-cover [km ²]
Zmutt	B57-05	2016	14.8	6.9	2995	19.4	36	3.4
Unteraar	A54g-11	2016	22.7	12.3	2649	18.4	82	7.1
Oberaletsch	B36-01	2017	17.1	9.5	2902	21.7	122	5.6

5 Methods

5.1 Developing the methodology

The methodology was developed through an iterative testing process until a final workflow was established, as sketched out in Figure 7. The whole process was implemented in Google Earth Engine (GEE) and the main results were visualised within GEE. Python, QGIS and Excel were used as additional tools for statistical analysis and data visualisation.

The workflow started with the selection of relevant data (Sections 5.1.1 and 5.1.2), being based on the Digital Elevation Model (DEM), satellite imagery from Landsat 5 and 8, and the glacier inventories introduced in the Data section. LST was normalised for elevation (Section 5.1.3) and a normalised difference index with NIR was calculated (Section 5.1.4). Various input layers were generated for the final classification, which are described in detail in Sections 5.1.5. The final classification was achieved with a Random Forest (RF) classifier and statistically evaluated (Section 5.1.6) against the ground reference derived from the glacier inventories (Section 5.1.7).

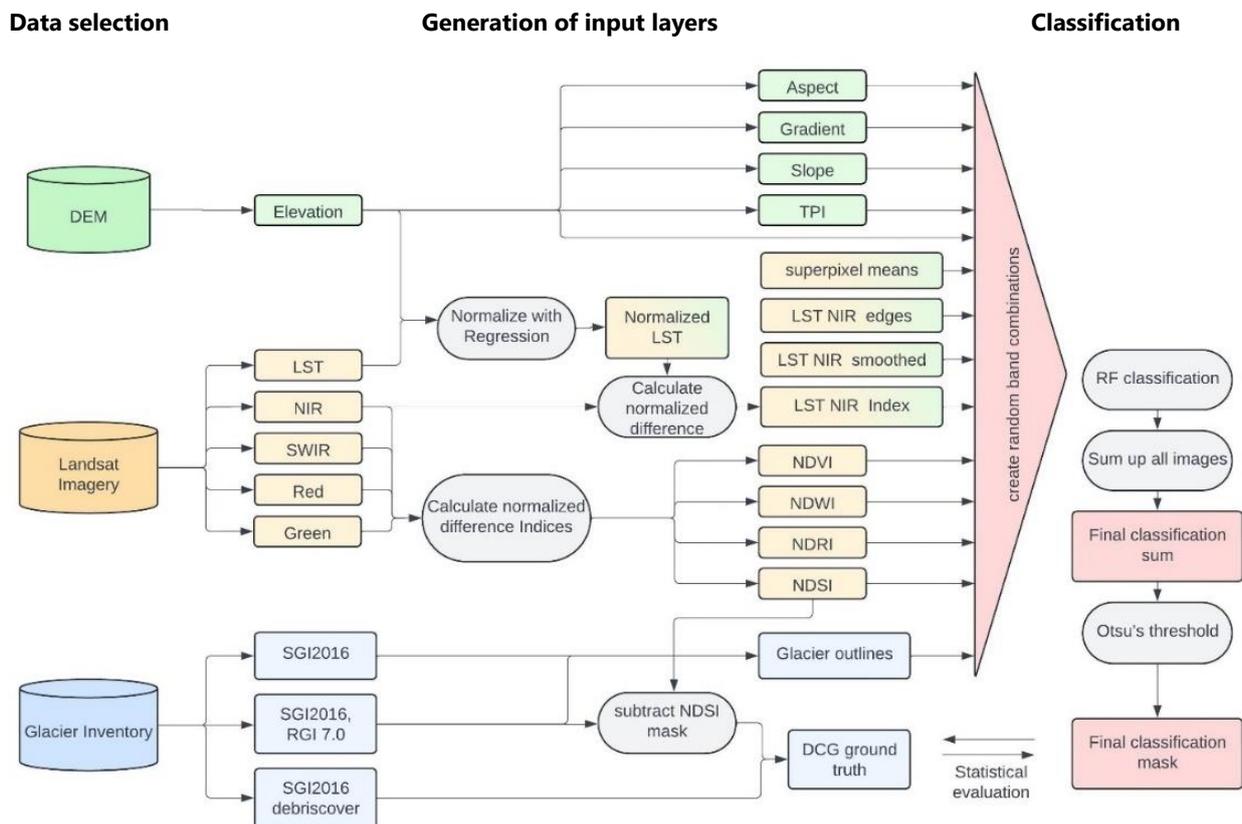


Figure 7: Flowchart of the process applied in GEE, starting with data selection of relevant input data for the generation of input layers that were used for DCG classification.

5.1.1 Manual image selection

In this study, individual scenes of Landsat imagery were selected to analyse glacier changes for the years 2003, 2010, and 2016, individually for Zmuttgletscher, Unteraargletscher, and Oberaletschgletscher. Landsat 5 data were used for 2003 and 2010, while Landsat 8 data were used for 2016. The decision to

use individual scenes, rather than more automated approaches, was motivated by the greater control over the data. Each selected scene was visually inspected to ensure a cloud-free image, where the glacier was not covered by shadows or snow. The use of individual scenes, rather than composite images, minimized the risk of artifacts that can occur in composites combining multiple images (Qiu et al., 2023). The use of individual, cloud-free scenes ensured that the data were consistent and reduced the likelihood of errors due to varying atmospheric conditions or artifacts from blending different scenes. This way, the performance of the method in identifying DCG could be better evaluated while eliminating as many other sources of error as possible.

An overview of the Landsat scenes that were selected is provided in Table 4. The years were chosen to align with the dates of the glacier inventories used for comparison and validation. Notably, the scenes used for Oberaletschgletscher for comparison to the SGI2016 is from 2016, although the acquisition date for Oberaletsch in the SGI2016 is from the year 2017. However, in summer 2017 no Landsat scene was found to show the whole glacier free of snow and cloud coverage, which is why a scene from summer 2016 was selected.

For Satopanth glacier, a scene from 2002 would have been ideal for comparability to the RGI 7.0. However, the closest scene available that showed the whole glacier without snow or cloud coverage was from the year 1998.

Table 4: Overview of the Landsat scenes selected for analysis of DCG in 2003, 2010 and 2016.

	Comparison to RGI 7.0 (2003)	Comparison to SGI2010 (2010)	Comparison to SGI2016 (2016)
Zmutt	LT05_195028_20030813	LT05_195028_20100901	LC08_195028_20160901
Unteraar	LT05_195028_20030813	LT05_194028_20100825	LC08_194028_20160825
Oberaletsch	LT05_195028_20030813	LT05_194028_20100825	LC08_194028_20160825
Belvedere	LT05_195028_20030813		
Satopanth	LT05_146039_19981011		

5.1.2 Pixel-wise selection of scene and composite creation

A second possibility for image selection is a more automated approach of pixel-wise selection of a scene and subsequent generation of a composite image. Landsat 5 and 8 images, respectively, were filtered to cover the period from June to October of the respective year to ensure comparability with the corresponding glacier inventory. The summer months were chosen to reduce the likelihood of snow cover, thus improving conditions for the observation of LST and reflections in the optical to infrared spectrum. To improve data quality, a cloud score algorithm was applied to remove pixels affected by cloud cover, leaving only clear sky observations. Cloud thresholds for the VIS, NIR, SWIR, and LST bands were used to calculate the score. To distinguish between clouds and snow, the Normalised Difference Snow Index (NDSI) was included:

$$NDSI = \frac{(Green - SWIR 1)}{(Green + SWIR 1)} \quad (1)$$

The pixel-wise selection was based on the LST value of each pixel. A higher LST value compared to the same pixel in other scenes reduces the probability of clouds, shadows, and snow-covered debris, as all these features would cause lower LST values (Kraaijenbrink et al., 2017). Instead of using maximum LST values, which can introduce outliers, the top 20% of LST values were selected to provide a more robust

representation of high temperature areas. These selected pixels were then aggregated into a median composite image, providing a stable and reliable dataset for subsequent analyses.

This approach could be used as an alternative first step in the workflow if the goal is to apply the methodology in a broader geographic context where visual inspection and manual image selection are too labour intensive. This approach was not applied for the generation of results in this thesis to isolate the performance of the methodology and reduce the influence of artifacts in the input data, as can be the case with composite images.

5.1.3 Normalisation of LST for elevation

To account for the influence of topographic variability on LST (Malbêteau et al., 2017), a normalisation process was employed to adjust LST for elevation effects. This normalisation was achieved through a systematic Approach Involving linear regression and subsequent rescaling.

Initially, a linear regression analysis was conducted to establish the relationship between elevation and LST, which enabled the extraction of the regression slope and intercept (Figure 8). This regression model served as a predictive tool for understanding how LST varies with changes in elevation.

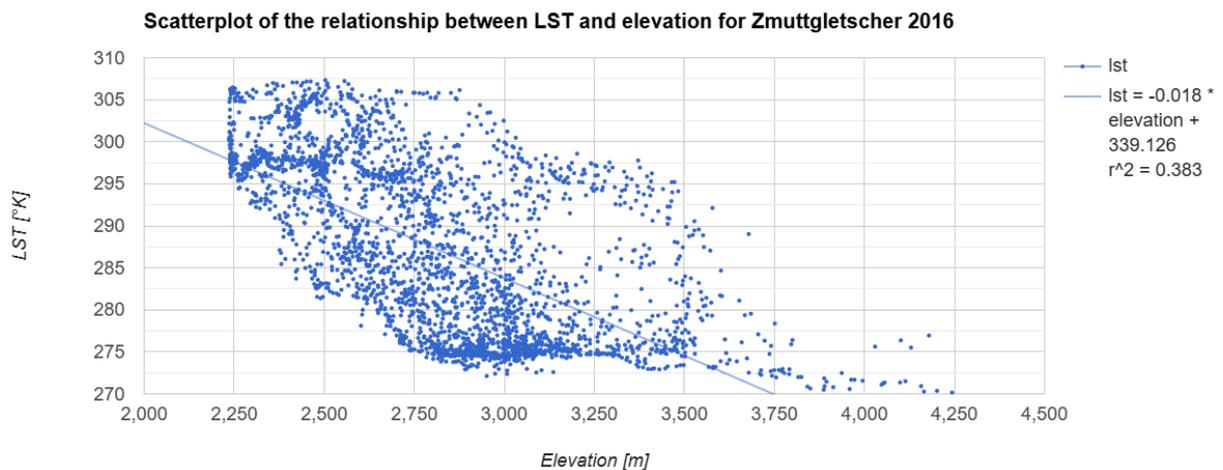


Figure 8: Scatterplot of the relationship between LST and elevation, exemplary for Zmuttgletscher 2016, showing 3000 points and including a trendline.

Using the established regression model, predicted LST values were calculated based on the elevation data. The original LST values were then normalised by subtracting these predicted values, resulting in a new set of adjusted LST data that accounts for elevation effects. The formula used can be described as:

$$LST_{normalised} = LST_{original} - LST_{predicted} \quad (2)$$

To further enhance the reliability of the normalised LST, a rescaling function was applied, standardizing the values to a range of [0, 1]. This function computed the mean and standard deviation of the normalised LST within a defined region of interest (ROI). Values that fell outside three standard deviations

from the mean were masked out to minimize the influence of outliers. The remaining data were then rescaled to ensure consistency and to further reduce the impact of elevation-induced variations.

The final output is a normalised and rescaled LST image with improved suitability for the subsequent classification steps, as shown in Figure 9 exemplary for Zmuttgletscher 2016.

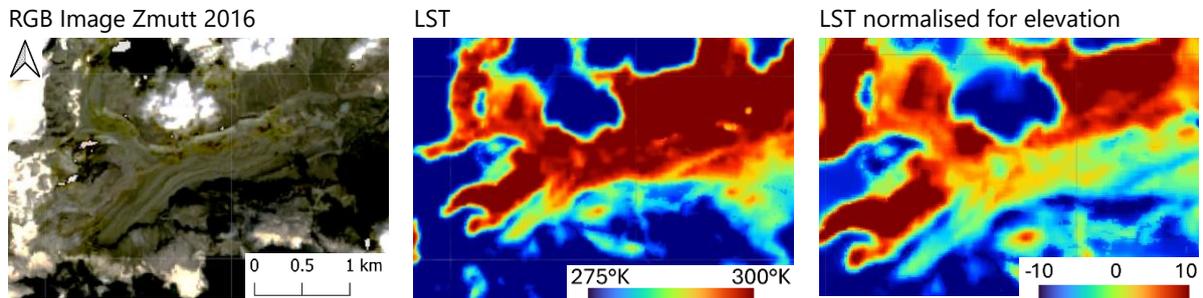


Figure 9: Normalisation of LST for elevation by regression, visualized for Zmuttgletscher 2016.

5.1.4 LST-NIR Index

The LST-NIR Index was calculated to enhance the detection of DCGs by combining the normalised LST and near-infrared (NIR) bands. The normalised LST (LST_{norm}) was integrated with NIR values to create a normalised difference index. This combination leverages the thermal properties captured by LST and the reflective characteristics of NIR to improve the identification of debris-covered areas. The formula for this index was:

$$LST\ NIR\ Index = \frac{LST_{norm} - NIR}{LST_{norm} + NIR} \quad (3)$$

To explore the value of the NIR band in combination with LST, a cross analysis between each band of the Landsat image and the normalised LST was conducted (overview in Figure 10). Combining the LST with any other band improved the differentiation between DCG and non-DCG on the visualized image, thanks to the addition of a second dimension. Out of the tested band combinations of LST with the blue, red, green, NIR, SWIR 1 and SWIR 2 bands, the combination with NIR improved visual differentiation between DCG and non-DCG the most.

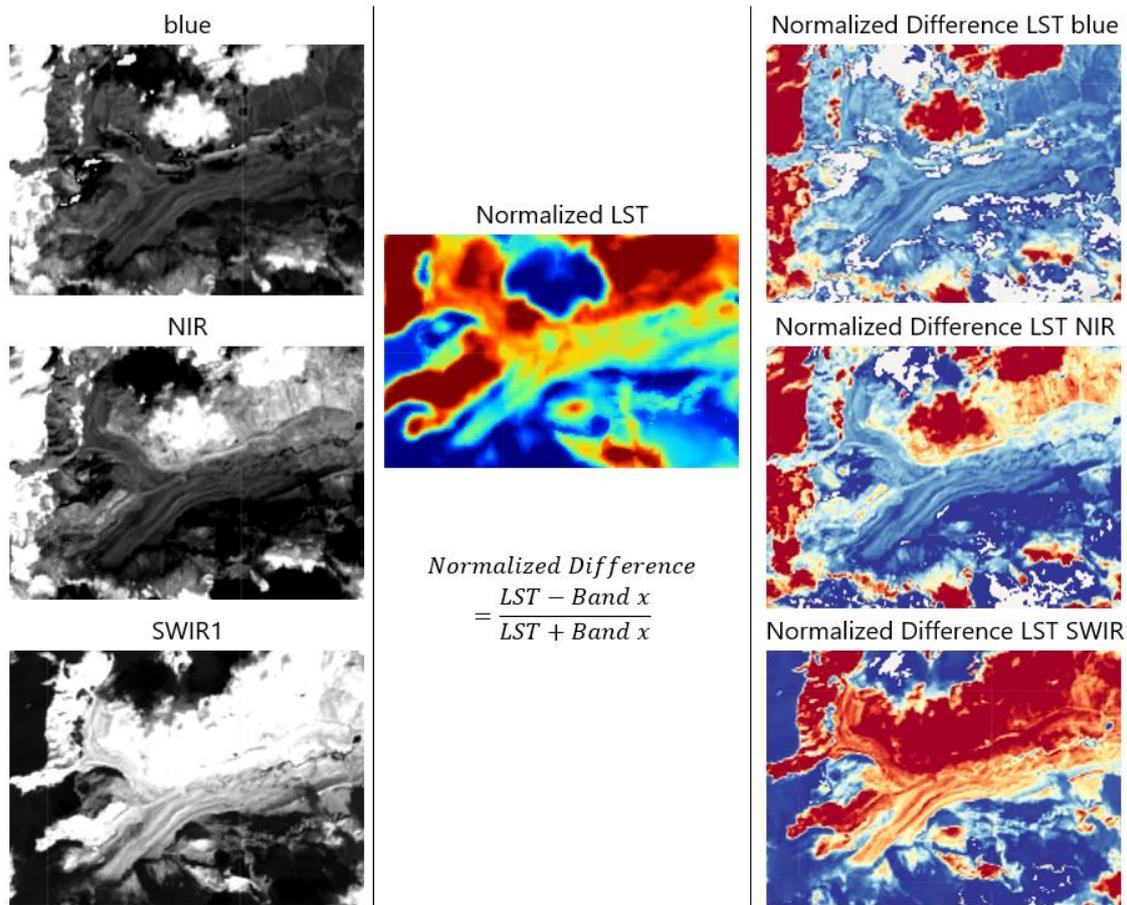


Figure 10: Visualisation of the comparison of combining different bands with LST to enhance differentiation between PGD and SGD, using equation (3) while substituting NIR with different bands ('Band x').

Additionally, the relationship between different spectral bands and LST was visualised using scatterplots, highlighting the benefits of incorporating a second dimension to distinguish DCG from non-DCG. This analysis was performed in Python, using 2016 Landsat imagery of Zmuttgletscher. The DCG and non-DCG classes used for this comparison were derived from the SGI2016 debris-cover class.

5.1.5 Classification with Random Forest

To classify DCGs, a Random Forest (RF) classifier was employed, utilizing a diverse set of input layers (Table 5). RF is an ensemble learning algorithm that constructs multiple decision trees during training and aggregates their predictions through majority voting to generate the final classification (Belgiu & Drăguț, 2016). This approach enhances accuracy and reduces the risk of overfitting compared to single decision trees, making RF well-suited for complex, high-dimensional datasets (Mellor et al., 2015).

Despite testing different configurations of tree numbers and sample sizes, these adjustments had a negligible impact on the overall classification performance. The final RF model was configured with 100 trees and a stratified sample size of 5,000 for training and 3,000 for testing. These settings, determined after testing different configurations, provided a robust balance between computational efficiency and classification accuracy.

The classification process included the following steps:

1. Input data

Twenty different datasets were created, each with a unique combination of input layers. The choice of input layers played a crucial role in influencing the classification outcomes, with their importance varying based on the year and the glacier from which the data was derived. To account for this variability and to streamline the classification process, 20 random band combinations were generated and used as inputs for the classifier.

2. Data sampling

To extract a sample from the training data, the 'stratifiedSample' function in GEE was used (GEE, 2024). This function extracts a specified number of random points from each class, ensuring a balanced sample across classes. The output is a feature collection, where each feature corresponds to a sampled point, containing one property for each band in the input image.

3. Initial classification

The RF classifier was applied to each dataset, producing 20 intermediate classification outputs where each pixel was labelled as either 0 (non-DCG) or 1 (DCG).

4. Summation

These 20 binary classifications were summed pixel by pixel to produce a classification sum image. The values in this image ranged from 0 to 20, indicating how many of the individual classifications identified a given pixel as DCG (0 = none, 20 = all).

5. Thresholding

Otsu's method was applied to the classification sum image to determine a threshold that allowed the final binary classification into DCG (1) and non-DCG (0). This technique is widely used in image processing for automatic binarization, determining the optimal threshold value based on the histogram's shape (Xu et al., 2011). Otsu's method assumes that the image is composed of two classes and calculates the threshold that minimizes the within-class variance for both, consequently maximizing the variance between the two classes, resulting in an optimal threshold for segmentation (Xu et al., 2011). This approach ensured that the two classes (DCG and non DCG) were optimally separated, providing an automated and effective solution for final DCG classification.

The final classification mask produced values of 0 and 1 representing non-DCG and DCG respectively. In a subsequent step, a smoothing filter was applied to the mask, followed by vectorisation, resulting in the delineation of DCG outlines.

Table 5: Overview of the input layers for the Random Forest classification.

Information content	Layers
Layers based on LST NIR Index	<ul style="list-style-type: none"> a. LST NIR Index b. Super pixel means (SNIC) c. Smoothing Filters d. Edge Detection
Topographic information	<ul style="list-style-type: none"> e. TPI (Topographic Positioning Index) f. Elevation g. Slope

	<ul style="list-style-type: none"> h. Gradient i. Aspect
Normalised Difference Indices	<ul style="list-style-type: none"> j. NDRI (Normalised Difference Rock Index) k. NDWI (Normalised Difference Water Index) l. NDSI (Normalised Difference Snow Index) m. NDVI (Normalised Difference Vegetation Index)
Glacier Inventory	<ul style="list-style-type: none"> n. Glacier outlines <i>Approach I: RGI 7.0 (2003), SGI2010 and SGI2016</i> <i>Approach II: RGI 7.0</i> <i>Approach III: no glacier outlines</i>

a. LST NIR Index

The resulting image of the Normalised Difference between LST and NIR bands served as a primary input to the classification.

b. Super-pixel Means (SNIC)

Derived using the Simple Non-Iterative Clustering (SNIC) algorithm, superpixel means represent clusters of pixels with similar characteristics, allowing the calculation of means for specific features, such as the LST NIR index, within each superpixel. This technique reduces noise and variability at the individual pixel level, improving the spatial coherence of the classification, and reduces the effect of outliers. To apply the SNIC algorithm to the LST-NIR image, a segment size of 5, a compactness of 1 and a connectivity of 4 were chosen.

In the SNIC algorithm, the segment size controls the approximate number of pixels in each superpixel, with a smaller size (such as 5) capturing finer local detail and larger sizes capturing broader patterns. Compactness controls the trade-off between colour similarity and spatial proximity, with a low value (1) resulting in more regular, spatially compact superpixels. A connectivity of 4 means that only pixels sharing an edge are considered neighbours, resulting in well-defined super-pixel boundaries. For each cluster generated by the algorithm, the mean LST-NIR value was selected.

c. Smoothing Filters

A smoothing filter was applied to the LST-NIR index using a boxcar kernel with a radius of 5 pixels in GEE. This process reduced local variation and noise by averaging pixel values within a defined neighbourhood, helping to smooth abrupt changes between adjacent pixels (GEE, 2023a). The result was a clearer representation of broader spatial patterns, while minimising finer local detail. This step helps to highlight general trends in the image and improves the robustness of subsequent analyses by reducing the influence of outliers or isolated noise.

d. Edge Detection

A 3x3 Sobel edge detection kernel was applied to the LST-NIR index to highlight areas of high spatial gradients, i.e. rapid changes between high and low values. This method emphasises the boundaries between different regions and helps to identify edges or transitions in the data, such as sharp changes

in land cover or temperature (Vincent & Folorunso, 2009). Default values of 1 were used for both magnitude and normalisation in the kernel. The result is a clearer delineation of boundaries in the LST-NIR index.

e. TPI (Topographic Position Index)

The Topographic Position Index (TPI) measures the relative elevation of a location compared to its surrounding terrain and is useful for distinguishing between valleys, slopes, and ridges (De Reu et al., 2013). The TPI was calculated by subtracting the mean elevation within a circular neighbourhood of 100 pixels from the elevation at each point using the following formula:

$$TPI = elevation - mean(elevation\ in\ neighbourhood) \quad (4)$$

The resulting TPI values highlight variations in topography. Positive TPI values indicate that a point is higher than its surroundings (e.g. ridges), while negative values suggest lower elevations (e.g. valleys). Values near zero indicate that the point is on relatively flat or evenly sloped terrain. The TPI values were derived by applying this method to the elevation data, highlighting variations in topography. This information is particularly useful in landscape classification, hydrological modelling, and understanding ecological patterns by identifying key landform features (De Reu et al., 2013).

f. Elevation

Elevation data from the digital elevation models (DEMs) were included in the classification as a separate input layer, although they are partially included in the LST-NIR index due to the normalisation of LST with respect to elevation. The decision to include elevation as a separate layer was driven by its significant influence on glacier distribution and its critical role in the classification process. Elevation directly affects temperature, ice formation and debris-cover, making it a key factor in distinguishing glacier characteristics, justifying its inclusion despite overlap with other input variables.

g. Slope

Slope is used to account for terrain influence on LST values, with low-slope DCGs and steep surrounding flanks serving as key features.

Visualizing the slope for several DCGs, i.e. Zmuttgletscher, Unteraargletscher, Oberaletsch-gletscher and Satopanth glacier, it could be observed that these glaciers all have a low slope, while being surrounded by mountain flanks of steep slopes. Other methodologies that were developed to identify or outline DCGs also included slope, mostly in the form of thresholds, below which a DCG is more likely (Nunchhani et al., 2024; Sharda & Srivastava, 2024).

The slope was derived from the DEM using the built-in function `ee.Terrain.slope` in GEE, which calculates the local gradient based on the four connected neighbours of each pixel (GEE, 2023e).

h. Gradient

The gradient describes the ratio between the elevation change of the slope and the horizontal length over which the change occurs. It captures the change in steepness, further informing surface and terrain variability. The gradient was calculated based on the DEM with the built-in function `image.gradient` in GEE. This function calculates the x and y gradients of an image, which were then used to calculate the magnitude according to the following equation, as suggested in the GEE documentation (GEE, 2023b).

$$\text{Gradient Magnitude} = \sqrt{(\text{x gradient})^2 + (\text{y gradient})^2} \quad (5)$$

i. Aspect

Aspect, or the direction of slope, helps identify directional effects on temperature and vegetation distribution. Aspect alone may not be informative for the presence and extent of a DCG, but in combination with other input layers that are based on LST or elevation, it adds a relevant piece of information. The aspect was calculated from the DEM, using the built-in function `ee.Terrain.aspect` in GEE (GEE, 2023d).

j. NDRI (Normalised Difference Rock Index)

The Normalised Difference Rock Index (NDRI) was calculated as the normalised difference between the SWIR 1 and the red band, as proposed by Huang & Cai (2009), to highlight debris-covered areas regardless of the presence of underlying glacier ice. To compute the NDRI, the built-in function `ee.Image.normalisedDifference` in GEE was used according to the following formula (GEE, 2023c):

$$\text{Normalized Difference} = \frac{(\text{Band 1} - \text{Band 2})}{(\text{Band 1} + \text{Band 2})} \quad (6)$$

where SWIR 1 was used as Band 1 and the red band as Band 2. This approach highlights areas where rock or debris is present, making it easier to distinguish debris-covered terrain from other surfaces.

k. NDWI (Normalised Difference Water Index)

The Normalised Difference Water Index (NDWI) was derived from the green and NIR bands to distinguish water bodies from debris-covered areas. The NDWI was calculated using the same normalised difference formula described previously for the NDRI, with the green band as Band 1 and the NIR band as Band 2. This index highlights water bodies, as water mostly absorbs NIR and reflects green light, resulting in higher NDWI values for water and lower values for land or vegetation (McFeeters, 1996).

l. NDSI (Normalised Difference Snow Index)

The Normalised Difference Snow Index (NDSI) was calculated using the green and SWIR 1 bands to distinguish snow and ice from debris-covered surfaces. The NDSI was calculated using the same normalised difference formula as previously described, with the green band as Band 1 and SWIR 1 as

Band 2. This index is particularly effective in identifying snow because snow strongly reflects green light and absorbs SWIR, resulting in higher NDSI values for snow and lower values for other surfaces (Hall et al., 1995).

m. NDVI (Normalised Difference Vegetation Index)

The Normalised Difference Vegetation Index (NDVI), calculated using the green (Band 1) and NIR (Band 2) bands, was employed to distinguish vegetated areas in the glacier forefield from debris-covered regions. The NDVI is effective in identifying vegetation, enabling the differentiation between vegetated and non-vegetated surfaces (Taloor et al., 2021). It is assumed that there is generally more vegetation on PGD than on SGD, based on the more stable and favourable conditions for plant growth in periglacial areas compared to the harsher, more dynamic environment of SGD (Sharda & Srivastava, 2024).

n. Glacier Outlines

Glacier outlines were obtained from the glacier inventory datasets RGI 7.0 (2003), SGI2010 (2010), and SGI2016 (2016). These outlines represent entire glacier extents but do not distinguish between debris-covered and clean ice areas.

In the primary approach, for each year, only the glacier outlines from the corresponding inventory were used as input layers, which limits the applicability of the method to years with current glacier outlines.

An alternative approach would be to use only the RGI 7.0 outlines for both training the classifier and for the testing region, even for years other than 2003. This allows the other input layers to provide updated information for the year of analysis, while the existing glacier outlines still serve as additional information for the classification. To assess different options, if and how to include glacier outlines in the classification, three approaches were established and tested:

- I. The glacier outlines from the corresponding year were used as an input layer. This requires a glacier outline of the year of interest for both the training and testing dataset, thus limiting the applicability of this approach to glaciers in Switzerland to the years 2003, 2010 and 2016, and on a global scale to the year 2003. This served as the primary approach to evaluate overall model performance.
- II. The 2003 glacier inventory (RGI 7.0) was used as an input layer consistently for both training and test datasets, regardless of the year of other input data.
- III. No glacier outlines were included. Instead of using the glacier inventory as an input layer, the number of band combinations was increased. This approach used 40 randomly generated band combinations, each consisting of 5 to 6 layers. Due to computational limitations, Otsu's threshold could not be calculated within the same script; therefore, thresholds were selected manually in this case to simplify processing.

The other Input layers (a-m) were used identically for all three approaches.

5.1.6 Statistical metrics for performance evaluation

To evaluate the performance of the RF classification, the results were not only visualised but also statistically evaluated using several metrics: accuracy, precision, recall and F1 score. These metrics are based on the number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) derived from the final classified image, providing a comprehensive analysis of the model's performance.

Accuracy measures the proportion of correctly classified instances out of the total instances. It is useful when the classes are balanced but can be misleading with imbalanced datasets. Precision is the proportion of true positive predictions among all positive predictions made by the model. It evaluates the accuracy of the positive predictions and is crucial in minimizing false positives. Recall measures the proportion of actual positives that were correctly identified by the model. It is important for minimizing false negatives and is particularly useful in identifying all relevant instances in a dataset. The F1-score is the harmonic mean of precision and recall, providing a single metric that balances their trade-off. It is particularly effective for evaluating models on imbalanced datasets. The following equations (7, 8, 9, 10) describe the calculation of these metrics (Haranadhdy & Karthikeyan, 2022).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$F1 \text{ score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (10)$$

5.1.7 Ground Reference

Statistical evaluation of the results of the methodology requires a ground reference for comparison. The RGI and SGI datasets have been used for this purpose, recognising that they do not represent absolute ground truth, but provide a widely accepted basis for validation.

For the year 2016, the SGI2016 dataset includes a specific "debris cover" class that delineates debris-covered glacier areas. This layer was used as a ground reference to test the results of the methodology when applied to the 2016 data.

The RGI 7.0 (2003) and SGI2010 datasets provide outlines of entire glacier areas but do not differentiate between DCG and non-DCG, unlike the SGI2016 dataset. To effectively assess the performance of the model in identifying DCGs in the years 2003 and 2010, it was essential to create a reliable ground reference dataset for comparison.

The following workflow was implemented to generate this ground reference for DCGs: First, the Normalised Difference Snow Index (NDSI) was calculated to distinguish between snow or ice, and debris-covered areas. A threshold value of 0.4 was applied to the NDSI to create a binary mask, where areas with NDSI values above this threshold were identified as snow or ice. This binary mask was multiplied with the glacier outlines to exclude regions with high NDSI values, thereby retaining only the areas corresponding to DCGs in the ground reference dataset. The resulting image was the glacier outline as

provided by either the SGI2010 or the RGI 7.0 but without areas that were identified as snow or ice. This approach ensured that the ground reference represents only DCG areas, allowing for a more robust evaluation of the classification performance.

5.2 Evaluation strategy

The methodology was evaluated through two complementary strategies: temporal analysis and glacier-to-glacier comparison. These strategies assessed the Random Forest (RF) classifier's ability to detect glacier changes over time and to generalize across different glaciers, using data from three available years (2003, 2010, and 2016) and multiple glacier sites.

5.2.1 Temporal change analysis

To assess the methodology's performance over time, the RF classifier was trained on data from a single glacier for two different years and tested on the third year. This analysis required multi-year ground reference data, which were available only for the Swiss glaciers: Zmuttgletscher, Unteraargletscher, and Oberaletschgletscher. Temporal evaluation was conducted for these glaciers using data from 2003, 2010, and 2016.

5.2.2 Cross-glacier analysis

To evaluate the classifier's ability to generalize across different glaciers, the RF model was trained on data of two different glaciers and tested on a third. This analysis included data from all three Swiss glaciers for the years 2003, 2010, and 2016.

Additionally, the methodology's potential for global application was assessed by applying the RF classifier trained on the three Swiss glaciers in 2003 to Belvedere Glacier (Italy) and Satopanth Glacier (Himalayas), using 2003 data for both. Performance was evaluated through statistical measures, including accuracy, precision, recall, and F1-score.

In the cross-glacier analysis, three approaches were tested to evaluate the impact of glacier outline selection on classification accuracy:

- Approach I: Glacier outlines corresponding to the year of analysis were included as input layers. This approach requires ground reference data for the year of analysis and was applied to all glaciers for 2003 (using RGI 7.0) and to the Swiss glaciers for 2010 (SGI2010) and 2016 (SGI2016).
- Approach II: Glacier outlines from 2003 (RGI 7.0) were used as input layers, regardless of the analysis year. This approach requires ground reference data from a year other than 2003 and was applied to the Swiss glaciers for 2010 and 2016.
- Approach III: No glacier outlines were included in the input layers. This approach requires ground reference data for the year of analysis and was applied to all glaciers for 2003, and to the Swiss glaciers for 2010 and 2016.

The approaches differ in their transferability to other glaciers and years. Although statistical validation is limited, Approaches II and III can be extended to any region and year of interest. The results from these three approaches provided insights into the sensitivity of the methodology to the inclusion or exclusion of glacier outlines, as well as its robustness when applied across different glacier environments.

6 Results

The results of the methodology are presented both statistically and visually, providing a comprehensive evaluation of the classifier's performance in identifying DCGs compared to the ground reference data. The statistical evaluation includes metrics such as validation accuracy, precision, recall and F1 score, as described in Section 5.1.6. These metrics quantify the ability of the classifier to correctly identify DCG and distinguish it from non-DCG areas.

The visualised results include a classification sum image and a final classification mask. The classification sum image represents the cumulative results of classifications made using different band combinations. Higher pixel values indicate areas that were consistently identified as DCG across multiple band combinations. The final classification mask is a binary output derived from the classification sum image by applying a threshold. It represents the final classification result, delineating DCG and non-DCG areas.

The methodology was applied to Zmuttgletscher, Unteraargletscher, Oberaletschgletscher, Belvedere, and Satopanth glacier. Statistical validation was limited by the availability of ground reference data, so not all approaches could be tested on each glacier (see Section 5.2.3 for specifications).

Validation accuracies across approaches were generally high, with values above 0.7, often exceeding 0.8, and nearly half achieving accuracies over 0.9 (Figure 11). Accuracy measures the proportion of correctly classified pixels but may not fully reflect the model's performance on imbalanced classes, such as DCG and non-DCG. Therefore, additional metrics like precision and recall were also evaluated.

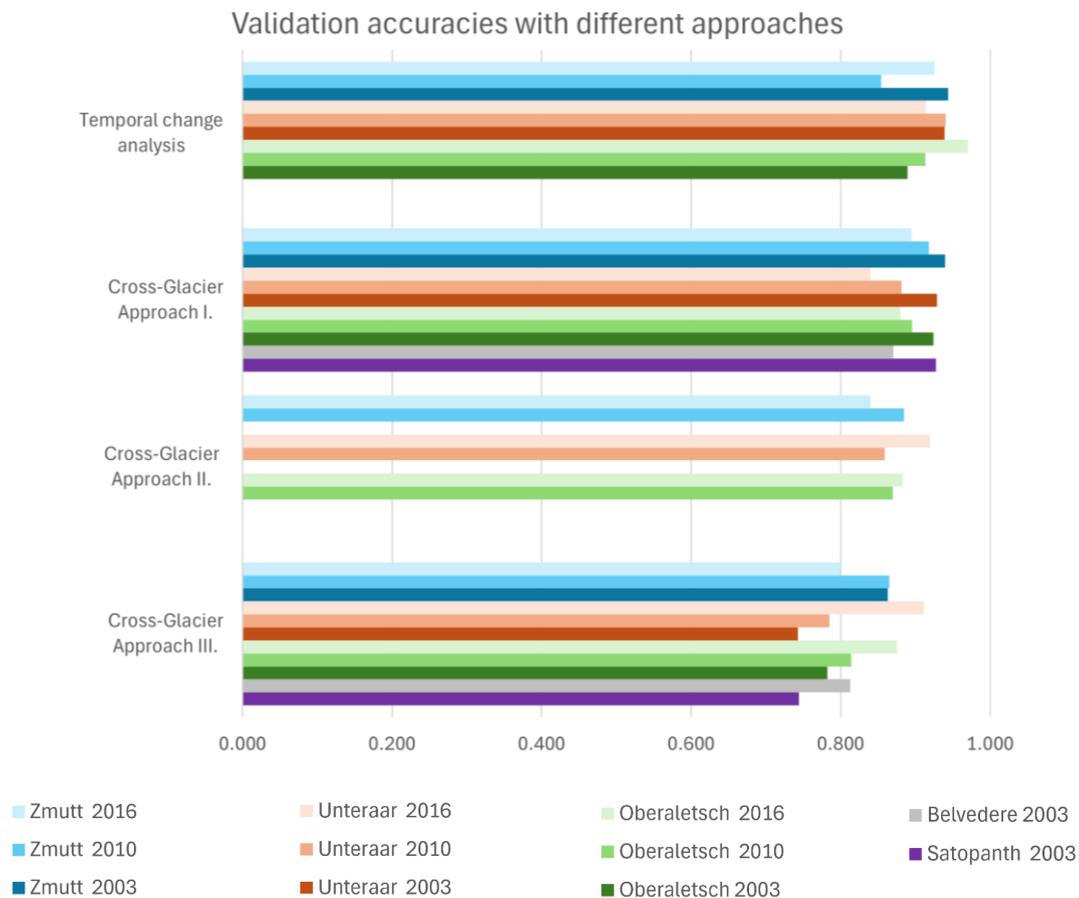


Figure 11: Overview of validation accuracies of classification results generated with different approaches, for all glaciers and years tested.

6.1 Temporal change analysis

The first step in evaluating the methodology was to apply it to one glacier at a time, with the aim of predicting glacier extent for one of the three years used in the analysis.

Due to variations in the input layers and glacier characteristics, the relative importance of different layers varied between years and glaciers analysed. To account for this, 20 random band combinations were generated and classified, and then merged into a single final classification using Otsu's threshold. Testing different numbers of band combinations and bands per combination showed that while increasing these numbers could improve the robustness of the approach, it generally did not lead to higher validation accuracies or improved quality of the mapped results for the glaciers analysed. The use of 20 band combinations with 5-6 bands per combination was found to be sufficient for primary analysis of the method.

Using different band combinations and combining them for the final classification was essential to capture different spectral characteristics across the datasets and to ensure consistent performance when applying the classifier across different glaciers and years. To compare the performance for different years and glaciers, Figure 12 shows the validation accuracies of the different band combinations sorted by their median accuracy. The final classification achieved the third highest median accuracy, with high accuracies (>0.85) for all glaciers and years tested. Of the two band combinations with higher median accuracies, only one achieved higher accuracies for all glaciers and years. Nevertheless, combining all accuracies into one final classification was preferred to using only this single high performing band combination, as the robustness of the method is expected to be greater with the combined approach. This is particularly important as testing the method for temporal changes was only a preliminary step before extending it to cross-glacier analysis.

Accuracies of the test year with multiple band combinations

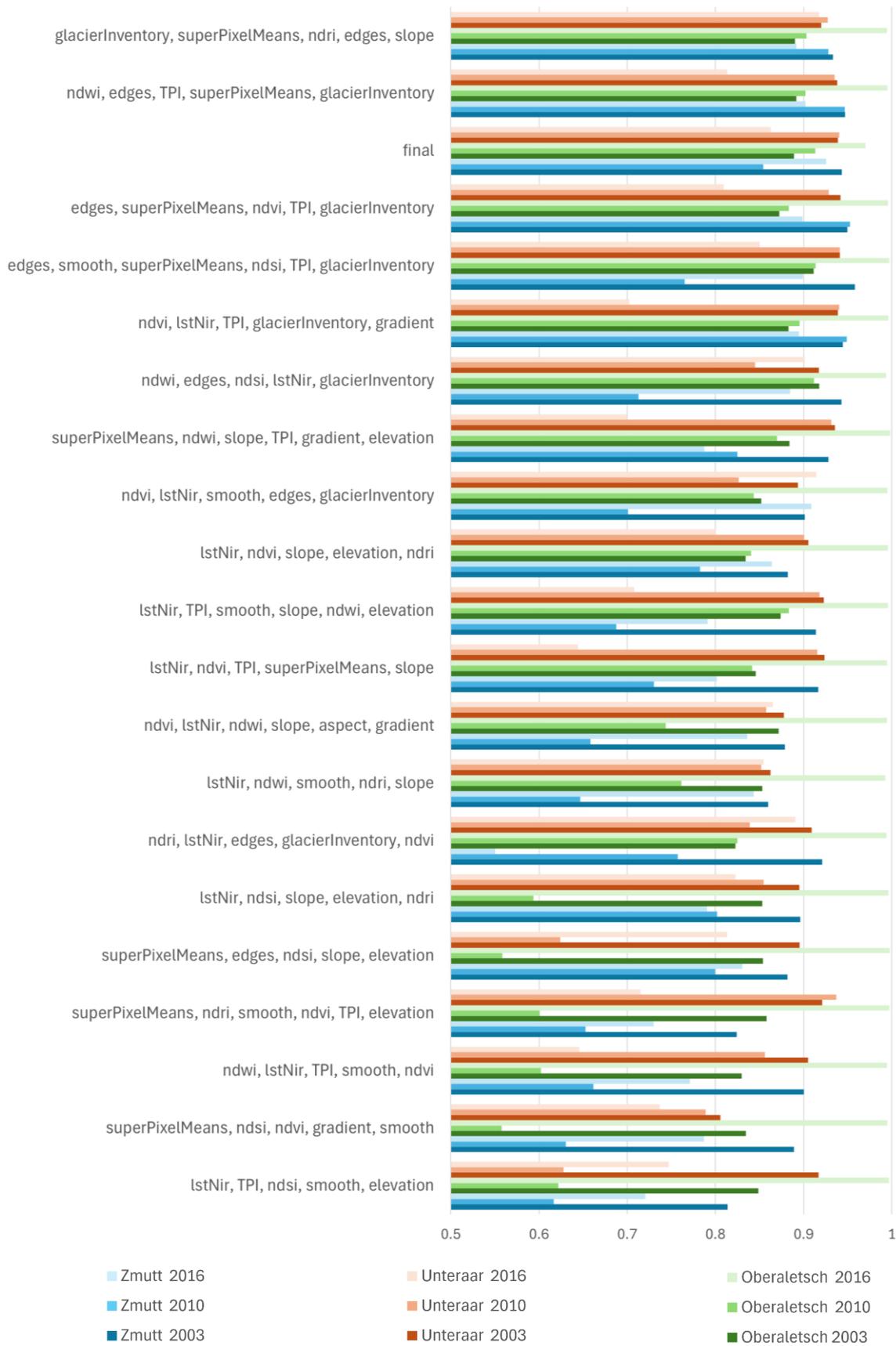


Figure 12: Accuracies of multiple band combinations used for the final classification, for the test year, when trained with the same glacier of two other years. The combinations are sorted by mean validation accuracy. The x-axis was adjusted to highlight differences in accuracy.

The Random Forest (RF) classifier was effective in identifying the main DCG areas when trained on data from two years and tested on a third year. The classifier’s performance was quantitatively assessed using statistical measures, including accuracy, precision, recall, and F1-score. These metrics were applied to evaluate the classification process, which ultimately produced a binary image representing the classified glacier change areas.

To optimize the binary classification, Otsu’s method was applied to determine an adaptive threshold for distinguishing between DCG and non-DCG. In comparison to a fixed, manually determined threshold (used uniformly across all images), Otsu’s method improved the validation accuracy in 4 out of 9 test cases (Figure 13). In these instances, the use of an adaptive threshold contributed to a more accurate delineation of DCG areas. In 3 cases, the validation accuracy remained unaffected, showing no improvement or decline when applying Otsu’s method. In 2 cases, the use of Otsu’s method resulted in a decrease in accuracy, suggesting that in some scenarios, a manually selected threshold may provide more consistent or reliable results.

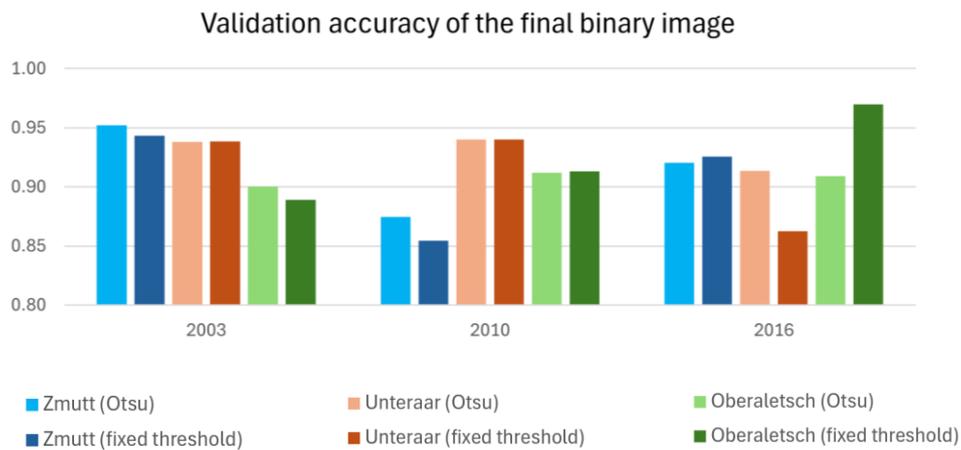


Figure 13: Comparison of validation accuracies per glacier and year with temporal change analysis, creating the final classification mask either with Otsu's threshold or with a manually selected, fixed threshold. The y-axis was adjusted to highlight differences in accuracy.

The statistical metrics used to assess model performance (Table 6) varied across glaciers and years. No systematic trends favouring either recall or precision were observed, indicating that the RF classifier neither systematically under- nor over-estimated DCG. The lowest accuracy was recorded for Zmuttgletscher in 2010, where a low recall indicated a significant portion of DCG pixels were missed. The corresponding high precision suggests that although fewer DCG areas were detected, they were classified accurately. This indicates a need for more careful threshold selection, potentially refining the adaptive method or further adjusting the fixed threshold.

Table 6: Overview of statistical metrics to evaluate the performance of the classifier for temporal changes, applied to Zmuttgletscher, Unteraargletscher and Oberaletschgletscher.

		2003	2010	2016
Zmutt	Accuracy	0.943	0.854	0.926
	F1-Score	0.945	0.842	0.923
	Precision	0.916	0.920	0.959
	Recall	0.976	0.776	0.889

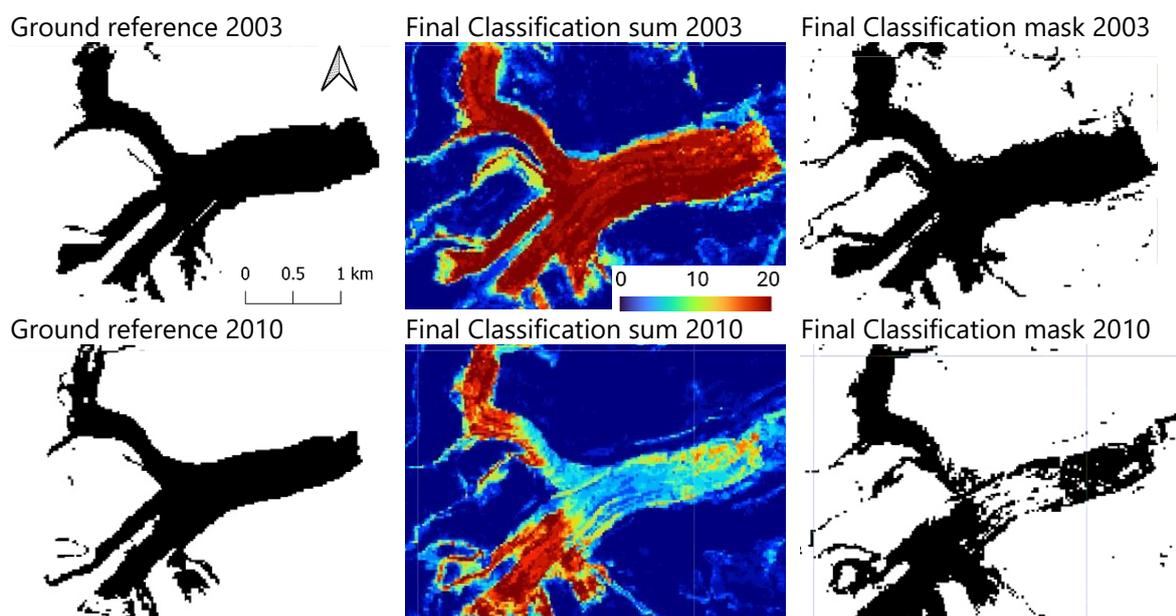
Unteraar	Accuracy	0.939	0.939	0.914
	F1-Score	0.939	0.939	0.906
	Precision	0.936	0.936	0.992
	Recall	0.942	0.942	0.834
Oberaletsch	Accuracy	0.889	0.913	0.970
	F1-Score	0.897	0.913	0.970
	Precision	0.838	0.918	0.960
	Recall	0.965	0.907	0.981

Zmuttgletscher

The mapped results for Zmuttgletscher, shown in Figure 14 show DCG delineation with high consistency compared to ground reference in 2003 and 2016, reflecting the high validation accuracies in these years. In 2010, the visualised results appear to be less accurate, as indicated by the slightly lower validation accuracy of 0.85, compared to accuracies greater than 0.93 in the other two years.

The final classification sum images of 2003 and 2016 both show high values for the DCG and low values for the surroundings. The thresholds chosen by Otsu's method result in a final classification mask that corresponds to the ground reference of the respective year, with only a few small outlying patches outside the glacier body being misclassified as DCG.

The final classification sum of Zmuttgletscher in 2010 shows a large area in the main glacier body that was not classified as DCG by the majority of classifications from different band combinations, resulting in low values in the sum image and a hole in the glacier in the final binary mask. As this area was clearly a debris-covered part of the glacier in 2010, this suggests an error in the classification. To identify the source of the error, the input layers of the training years (2003 and 2016) and the test year (2010) were visualised and compared. In this comparison, neither the pattern of the value distribution nor the value ranges varied significantly between the data from the training and test years. Overall, there is variation in the input data, but it is not restricted to the area in the middle of the glacier, nor does it indicate that 2010 is a general outlier in the comparison.



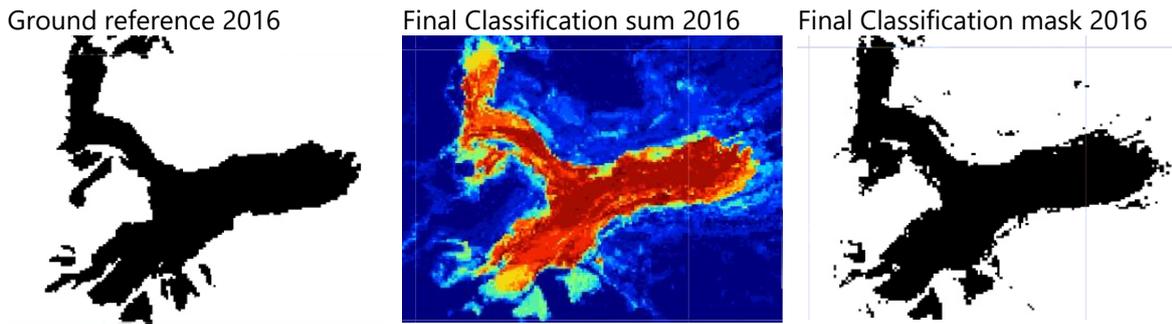


Figure 14: Classification results of Temporal changes, Zmuttgletscher. The colorbar in the final classification sum shows the number of times a pixel was classified as DCG by the different band combinations.

Unteraargletscher

The visualised classification results of Unteraargletscher (Appendix, Figure B 1) show a strong agreement with the ground reference, which is consistent with the high accuracy values observed in the statistical evaluation. However, closer examination of the classification sum and the final classification mask suggests that a more precise selection of thresholds could potentially yield better results.

In 2003, the highest values in the classification sum image are close to the ground reference, but the final classification mask appears to have a threshold set too low, resulting in an overestimation of DCG at the glacier terminus compared to the ground reference. A similar pattern emerges in 2010, where areas of high values in the classification sum again correspond more closely to the ground reference than to the binary classification mask. In both years, the classifier tends to extend the glacier slightly beyond the ground reference outlines.

In contrast, in 2016, both the classification sum and the final classification mask show a very high correspondence with the ground reference, indicating improved threshold selection and overall performance for that year.

Oberaletschgletscher

Similar to the patterns observed for Unteraargletscher, the final classification sum images for Oberaletschgletscher closely reflect the DCG delineation of the ground reference (Figure B 2).

In 2003 however, the final classification mask seems to include too many pixels along the whole glacier outline, indicating again a need for a lower threshold to generate the binary mask. Visual inspection of the areas of high values in the classification sum image show rather clear delineation of the DCG, corresponding well with the ground reference.

In 2010, the final classification mask aligns well with the ground reference, showing only some distinct patches of presumably wrongly classified DCG pixels outside the main glacier area.

In 2016, the final classification sum shows a distinctive transition from higher to lower values on the glacier tongue, that seems more like an artefact than a natural feature. Nevertheless, the values on the glacier tongue are still high enough compared to the surrounding to be classified as DCG in the final mask. Consequently, the result corresponds well with the ground reference.

6.2 Cross-glacier analysis

Cross-glacier analysis refers to the classifier being trained with data of one year from two different glaciers (when applied on Swiss glaciers) respectively three glaciers (when extended to Belvedere and Satopanth glaciers) and being tested on a third, respectively fourth glacier. Considering existing glacier outlines as an input layer, three different approaches were tested.

6.2.1 Approach I

The results of the statistical evaluation of the final classification (Table 7) show an overall high accuracy. For all three glaciers, the highest accuracies were achieved in 2003. Since the classifier for each tested glacier and year was trained with the two other glaciers of the respective year, this may indicate a higher consistency in the data between the different ROIs in 2003 than in the years 2010 and 2016.

Precision and recall in most cases lie relatively close together, indicating the threshold applied to the final classification sum is effectively distinguishing between the two classes DCG and non-DCG. All Otsu thresholds lie between 6 and 9, indicating that out of the 20 individual classification results from different band combinations, less than half produced accurate DCG classification on an individual pixel level.

Table 7: Overview of statistical metrics to evaluate the performance of the classifier for cross-glacier analysis, Approach I, applied to Zmuttgletscher, Unteraargletscher and Oberaletschgletscher in 2003, 2010 and 2016, and to Belvedere and Satopanth glaciers in 2003.

		2003	2010	2016
Zmutt	Accuracy	0.903	0.881	0.909
	F1-Score	0.906	0.890	0.911
	Precision	0.880	0.828	0.883
	Recall	0.934	0.962	0.942
	Otsu	9.000	8.000	7.000
Unteraar	Accuracy	0.868	0.888	0.950
	F1-Score	0.860	0.886	0.951
	Precision	0.913	0.904	0.937
	Recall	0.813	0.869	0.965
	Otsu	7.000	7.000	8.000
Oberaletsch	Accuracy	0.854	0.894	0.924
	F1-Score	0.860	0.895	0.918
	Precision	0.825	0.888	0.987
	Recall	0.898	0.902	0.859
	Otsu	8.000	8.000	8.000
Belvedere	Accuracy	0.866		
	F1-Score	0.873		
	Precision	0.830		
	Recall	0.922		
	Otsu	7.000		
Satopanth	Accuracy	0.941		
	F1-Score	0.944		
	Precision	0.899		
	Recall	0.995		
	Otsu	6.000		

Zmuttgletscher

Overall, the final classification sum and mask images for each year for Zmuttgletscher correspond well to the ground reference for that year (Figure 15). One feature that can be observed in all three sum images is a part of the glacier front that shows significantly lower values than the main glacier body. According to the ground reference, these areas should be classified as DCG, which was done correctly in the final classification mask for all three glaciers, indicating a good performance of Otsu's threshold.

The final classification of Zmuttgletscher 2010 shows a cohesive glacier body, in contrast to the result for Zmuttgletscher 2010 with temporal change analysis (section 6.1, Figure 14), where the classifier was trained with the two other years instead of two other glaciers.

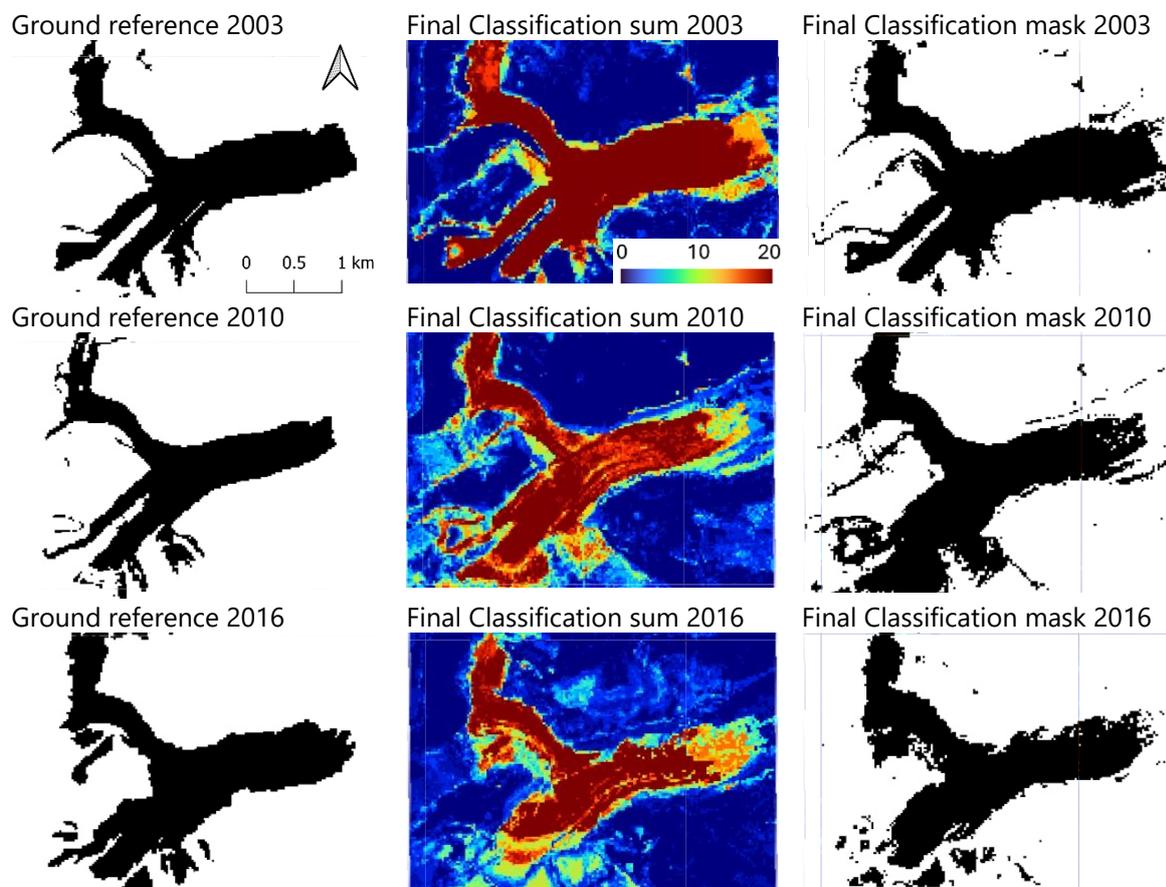


Figure 15: Classification results of Cross-glacier analysis, Approach I, Zmuttgletscher. The colorbar in the final classification sum shows the number of times a pixel was classified as DCG by the different band combinations.

Unteraargletscher

The results for Unteraargletscher with Approach I are found in the Appendix, Figure B 3. In 2003, the results show good agreement between the final classification sum and the final classification mask with the ground reference. However, a small area at the glacier front was classified as DCG, whereas it is not according to ground reference. In the sum image, this area has lower values than the rest of the glacier body, so applying a higher threshold would have resulted in a final mask that more closely resembled the ground reference. However, lowering the threshold would reduce the accuracy in the upper glacier arms, where similarly low values are correctly classified as DCG.

Both in 2010 and 2016, the high values in the final classification sum images align well with the corresponding ground reference. In both cases, the alignment slightly decreases in the final classification mask, with too many pixels at the glacier front and its side being, as well as some outlying patches, falsely classified as DCG. This suggests again a better result could have been achieved by applying a more restrictive threshold.

Oberaletschgletscher

Having been trained on Unteraargletscher and Zmuttgletscher, the classifier produced fairly accurate results when tested on Oberaletschgletscher, with final classification masks closely resembling the corresponding ground reference and only little to no outlying patches or 'noise' (Figure B 4).

In both 2003 and 2010, the areas with the highest values in the final classification sum images correspond well with the respective ground reference. However, in both cases, it appears that a slightly higher threshold might result in a more accurate final classification mask, as the current threshold seems to include too many pixels as DCG along the entire glacier outline.

In 2016, the values on the glacier are less uniformly high, but the threshold appears to be accurately chosen, effectively discriminating DCG from non-DCG in the final classification mask.

Belvedere Glacier

For Belvedere Glacier, the same approach was used, with the classifier trained on all three Swiss glaciers and tested on Belvedere. As only the RGI 7.0 is available to generate the ground reference for Belvedere, the statistical evaluation of the results is limited to the year 2003. In that year, the classification of Belvedere Glacier achieved a validation accuracy of 0.866, indicating a satisfactory performance of the method even when applied to a geographically distinct glacier. With a precision of 0.830 and a recall of 0.922, the threshold of 7, derived from Otsu's method, seems to efficiently discriminate between DCG and non-DCG.

In the visualised final classification sum, the main DCG body shows the highest values, although there are some outlying patches with high but slightly lower values (Figure B 5). In the final classification mask, many of these patches were misclassified as DCG. However, lowering the threshold would result in missing DCG pixels along the eastern outline of the DCG body. In addition, the final classification shows two medium-sized holes in the DCG body. While one of these is clearly visible in the ground reference, the other is only suggested there and is significantly larger in the classification result.

Satopanth Glacier

Extending the methodology used for Belvedere Glacier to Satopanth proved difficult due to the limited availability of Landsat 5 imagery for Satopanth. The RGI 7.0 outlines for Satopanth Glacier are based on 2002 imagery, so finding a Landsat 5 scene close to that year would be ideal for data consistency. However, no Landsat 5 imagery was available for Satopanth between 2002 and 2007. Extending the time span to include images from 1999 to 2008 resulted in 24 scenes, but none passed visual inspection due to cloud cover, snow cover, shadows or incomplete coverage of the region of interest (ROI). The closest suitable scene identified was from 11th of October 1998 (Figure 16), which captured the entire glacier without snow, clouds, or shadow obstruction.

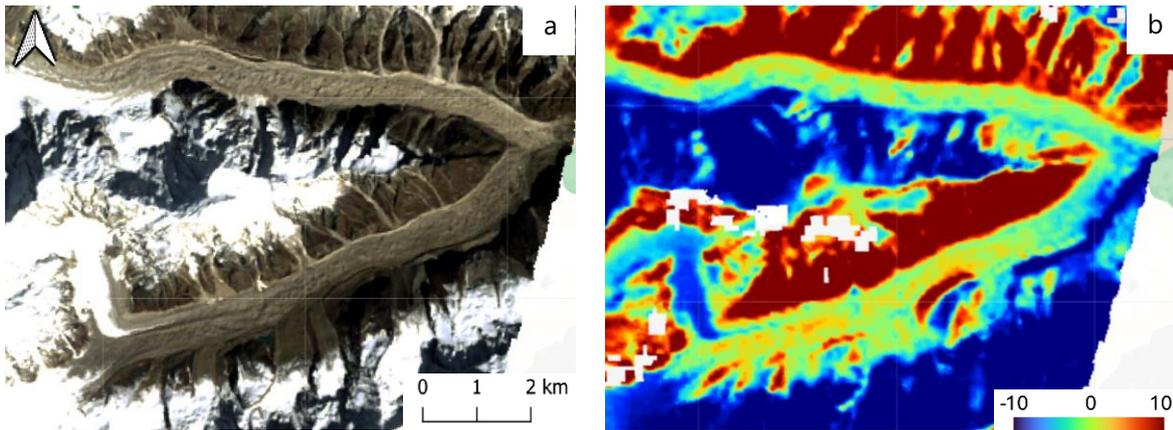


Figure 16: RGB image (a) and normalised LST (b) of the selected scene from 11th of October 1998 for the analysis of Satopanth glacier.

With a validation accuracy of 0.941, the classifier showed strong performance on Satopanth Glacier. The visualised results closely match the ground reference, with medium to high values in the classification sum image across the main DCG area (Figure B 6). Distinct transitions to surrounding areas with very low values further facilitate the distinction between DCG and non-DCG. The final classification mask corresponds well to the ground reference for DCG in 2003, although minor noise around the glacier margins slightly reduces the clarity of the delineation. Applying a smoothing filter prior to vectorising the results could address this issue. In addition, an elongated feature of misclassified DCG appears to the south-east of the main glacier tongue.

6.2.2 Approach II

Approach II, using only the 2003 glacier inventory (RGI 7.0), was tested on Zmuttgletscher, Unteraargletscher and Oberaletschgletscher for the years 2010 and 2016. In this setup, all other input layers were identical to the ones used in Approach I, selecting satellite images from summer 2010, respectively 2016 and using Copernicus data for 2010 and SwissAlti3D data for 2016. Applying this approach to the year 2003 would equal Approach I and was thus not included here.

The validation accuracies (Table 8) are generally high, indicating good classifier performance with this approach. The final thresholds for the classification mask, set using Otsu's method, range from 7 to 9, comparable to that of Approach I. This suggests a similar number of band combinations that result in effective DCG discriminations. Precision and recall values are closely aligned, confirming that the chosen threshold effectively discriminates between DCG and non-DCG.

However, the visualised results for Zmuttgletscher show a relatively poor delineation of the DCG body (Figure 17). Several patches outside the DCG are misclassified as DCG, and the glacier tongue appears longer than in ground reference, closely resembling the 2003 outline. In contrast, the visualised classification results for Unteraargletscher (Appendix, Figure B 7) and Oberaletschgletscher (Figure B 8) closely match the ground reference and effectively capture the DCG outlines.

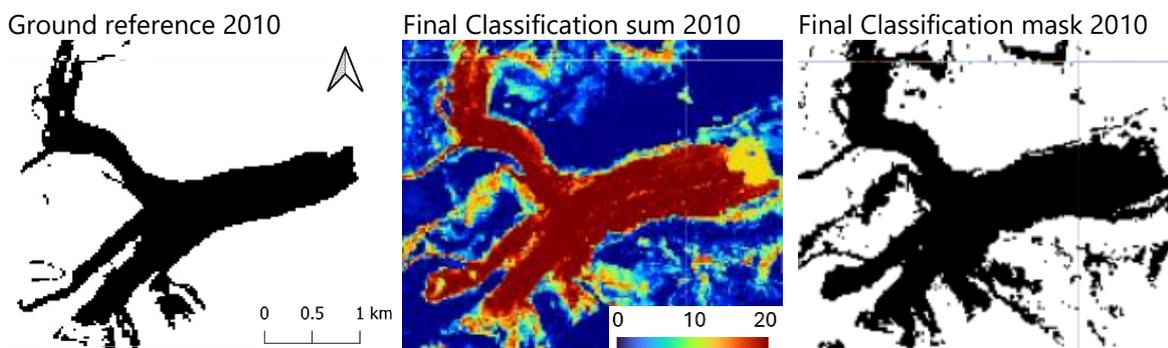
Table 8: Overview of statistical metrics to evaluate the performance of the classifier for cross-glacier analysis, Approach II, applied to Zmuttgletscher, Unteraargletscher and Oberaletsch in 2010 and 2016.

		2010	2016
Zmutt	Accuracy:	0.885	0.840
	F1-Score:	0.892	0.844
	Precision:	0.838	0.823
	Recall:	0.954	0.866
	Otsu:	9.000	8.000
Unteraar	Accuracy:	0.859	0.919
	F1-Score:	0.854	0.920
	Precision:	0.886	0.908
	Recall:	0.824	0.933
	Otsu:	7.000	8.000
Oberaletsch	Accuracy:	0.870	0.882
	F1-Score:	0.870	0.872
	Precision:	0.871	0.958
	Recall:	0.868	0.800
	Otsu:	8.000	8.000

Zmuttgletscher

While the main glacier body of Zmuttgletscher with Approach II was identified (Figure 17), the delineation of its outlines is noticeably less precise compared to the previous approaches (temporal change analysis, Figure 14 and cross-glacier analysis Approach I, Figure 15). For both the years 2010 and 2016, several patches outside the DCG are misclassified as DCG, and the glacier tongue appears longer than in ground reference, closely resembling the 2003 outline.

In 2016, the main DCG area has medium values in the final classification sum. To ensure this region was correctly classified as DCG, a lower threshold was set for the final classification mask. However, this lower threshold resulted in extensive outlying patches being incorrectly identified as DCG, such as those south of the main glacier tongue. Additionally, the glacier tongue appears overly extended compared to the ground reference, indicating an overestimation of the glacier area.



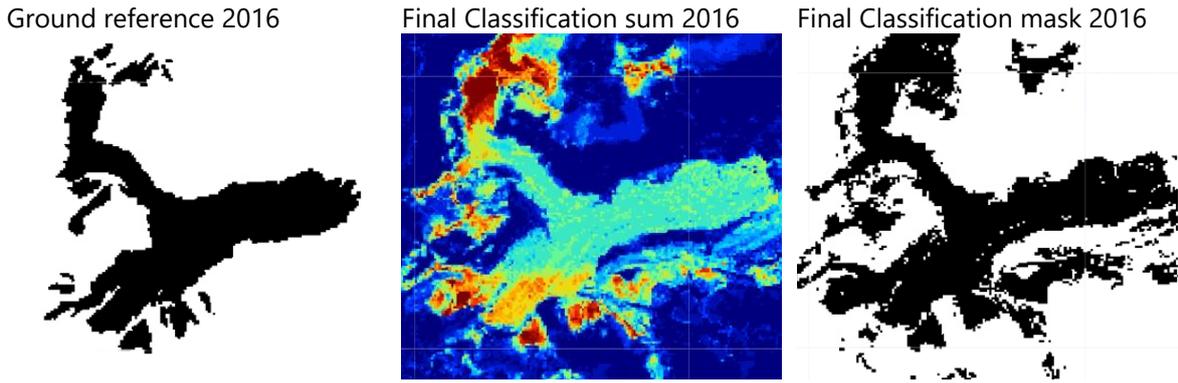


Figure 17: Classification results of Cross-glacier analysis, Approach II, Zmuttgletscher. The colorbar in the final classification sum shows the number of times a pixel was classified as DCG by the different band combinations.

Unteraargletscher

For Unteraargletscher, the results align more closely with the ground reference (Figure B 7) than those for Zmuttgletscher. However, in both 2010 and 2016, higher values in the classification sum image match the ground reference outlines more precisely than the final classification mask, suggesting that a higher threshold could improve accuracy. Nevertheless, the classifier worked effectively here. The temporal comparison clearly shows the retreat of the glacier terminus from 2010 to 2016, evident in both the sum and mask images. In addition, the lake in front of Unteraargletscher is no longer misclassified as DCG in 2016, resolving a problem previously seen with Approach I.

In 2010, the main glacier body aligns well with the ground reference, except for a slightly extended glacier front. In the final classification sum, high values closely match the ground reference terminus, though the final classification mask incorrectly includes additional pixels along the lake in front of the glacier as DCG. Outside the main glacier body, some round-shaped patches were misclassified as DCG, likely reducing validation accuracy; however, these can mostly be removed with a smoothing filter.

With a high validation accuracy of 0.919 in 2016 (Table 8), the classifier effectively captured the main DCG body and produced a fairly accurate delineation of the glacier front (Figure B 7). Misclassifications occurred along the southern edge of the main glacier body, where several pixel branches were incorrectly classified as DCG. Although the glacier terminus is relatively accurate in the classification sum, the final classification mask includes extra pixels, extending the terminus beyond the ground reference. Additionally, a line of pixels in the lake in front of the glacier was misclassified as DCG.

Oberaletschgletscher

The 2003 classification of Oberaletschgletscher shows numerous outlying patches misclassified as DCG (Figure B 8). While the main glacier body aligns well with the ground reference, the glacier front appears slightly narrower and less compact in the final classification mask.

The 2016 final classification mask demonstrates strong alignment with the ground reference. A notable feature is the distinct transition from high to medium values across the main DCG body in the classification sum image, though visual inspection of the input layers did not reveal a cause for this transition. The glacier front again appears less compact in the final classification mask, but this time reflecting the narrower delineation seen in the ground reference for this area in 2016.

6.2.3 Approach III

The approach of not using glacier outlines as input layers in either the training or test datasets was tested on Zmuttgletscher, Unteraargletscher and Oberaletschgletscher in 2003, 2010 and 2016 and on Belvedere and Satopanth glacier in 2003. Again, satellite images from the respective year and different DEMs were used to keep all other input layers identical to the first approach.

Due to computation efficiency, the threshold could no longer be determined using Otsu's method within the same script, as the number of band combinations was increased to 40. For reasons of simplicity, the thresholds were selected manually based on visual inspection of the final classification sum image.

The validation accuracies (Table 9) are slightly lower than in Approaches I and II, but still satisfactory, especially for Unteraargletscher and Oberaletschgletscher. Zmuttgletscher, with a validation accuracy of 0.799, again scores lower than the other two glaciers. The mapped results also show similar patterns as in the previous approaches, although this time the outline delineation for Zmuttgletscher is even weaker, failing to consistently capture the main glacier body and misclassifying large areas outside the glacier body as DCG (Figure 18). For Unteraargletscher (Figure B 9) and Oberaletschgletscher (Figure B 10), the classifier works well in identifying the main DCG body and delineates the glacier termini with satisfactory accuracy.

Table 9: Overview of statistical metrics to evaluate the performance of the classifier for cross-glacier analysis, Approach III, applied to Zmuttgletscher, Unteraargletscher and Oberaletschgletscher in 2003, 2010 and 2016, and to Belvedere and Satopanth glaciers in 2003.

		2003	2010	2016
Zmutt	Accuracy	0.863	0.865	0.799
	F1-Score	0.865	0.872	0.791
	Precision	0.848	0.825	0.825
	Recall	0.883	0.926	0.760
	Threshold	20	20	10
Unteraar	Accuracy	0.743	0.785	0.911
	F1-Score	0.653	0.741	0.910
	Precision	0.939	0.934	0.925
	Recall	0.501	0.610	0.894
	Threshold	20	25	25
Oberaletsch	Accuracy	0.823	0.826	0.881
	F1-Score	0.807	0.807	0.871
	Precision	0.889	0.905	0.950
	Recall	0.739	0.729	0.805
	Threshold	35	25	15
Belvedere	Accuracy	0.792		
	F1-Score	0.791		
	Precision	0.797		
	Recall	0.785		
	Threshold	20		
Satopanth	Accuracy	0.744		
	F1-Score	0.766		
	Precision	0.706		
	Recall	0.838		
	Threshold	5		

Zmuttgletscher

Approach III classification sum images for Zmuttgletscher in 2003 and 2010 show high values across the main DCG body (Figure 18), suggesting satisfactory DCG identification using multiple band combinations. The glacier fronts in these years show a clear transition from high to medium-low values but appear wider than the ground reference.

In contrast, the 2016 image shows low to medium values throughout the main glacier body, providing limited discrimination between DCG and surrounding areas (Figure 18). The reduced validation accuracy of 0.799 in 2016 (Table 9) corresponds to significantly weaker DCG identification in the classification output, with extensive misclassified outlying areas and missing sections in the main glacier body. All three results include a significant number of outlying areas misclassified as DCG.

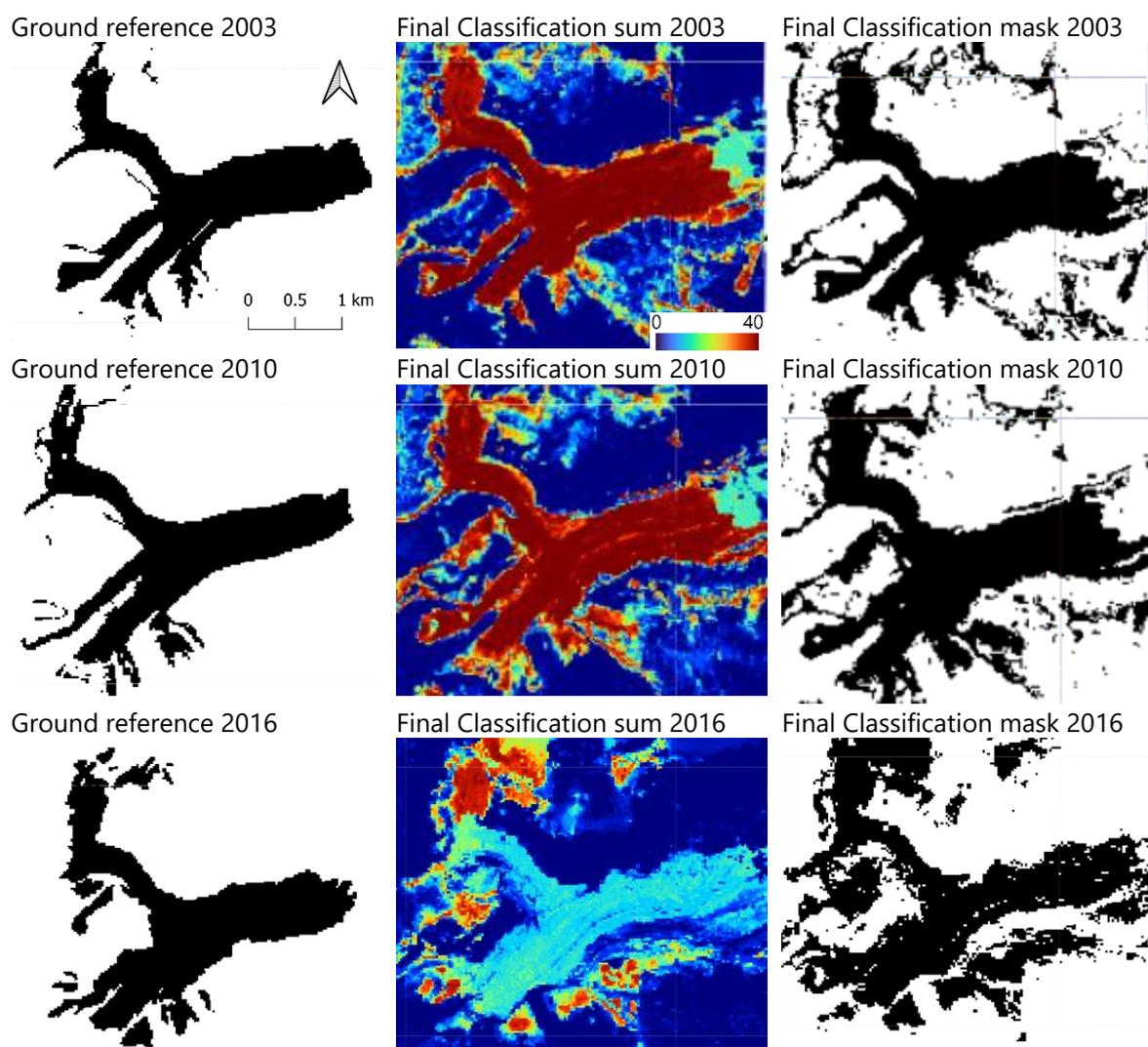


Figure 18: Classification results of Cross-glacier analysis, Approach III, Zmuttgletscher. The colorbar in the final classification sum shows the number of band combinations.

Unteraargletscher

Although the validation accuracies with Approach III in 2003 and 2010 are lower than for the previously introduced results for Zmuttgletscher with Approach III (Table 9), the visualised results show a clearer

outline delineation for Unteraargletscher (Figure B 9). In all three years, the classifier effectively delineated the main body of DCG, closely matching the ground reference outlines. The glacier fronts in 2003 and 2010 appear slightly extended compared to the ground references, but accurately capture the retreat of the glacier until 2016.

Oberaletschgletscher

Oberaletschgletscher achieves higher validation accuracies than Zmuttgletscher and Unteraargletscher when analysed using Approach III (Table 9). The visualized results of Oberaletschgletscher (Figure B 10) are comparable to those of Unteraargletscher (Figure B 9) in delineating the glaciers, with a generally high correspondence to the respective ground reference but also some misclassifications of both DCG and non-DCG.

The final classification sum for 2003 shows high values across the main DCG, with well-defined transitions to the surrounding areas. Various outlying features exhibit medium to medium-high values but are correctly excluded from the final classification mask. This suggests that the chosen threshold effectively distinguishes DCG from non-DCG, contributing to the overall accuracy of the classification.

In 2010, while the DCG is predominantly characterised by high values, the glacier tongue shows greater variability in values, with some lower values interspersed within the high value DCG. Additionally, a significant number of outlying non-DCG features display medium-high values, leading to a threshold for the final classification mask that fails to perfectly separate DCG from non-DCG. This results in the inclusion of some outlying patches and noise, as well as false negatives on the glacier tongue. These false negatives are particularly pronounced at the glacier terminus, where the DCG is not clearly delineated.

In 2016, the classification sum image is characterised by a distinct value transition along the main DCG tongue, similar to the pattern observed with Approach II. The main DCG has very high values that closely match the outlines of the ground reference. A sharp boundary marks the transition to the frontal ~1 km of the glacier tongue, where values are significantly lower. As Approach III involved manual threshold selection, the threshold was set low enough to include this section as part of the DCG area in the final classification mask, thus mitigating the impact of this pronounced feature.

Belvedere glacier

When applying Approach III to Belvedere Glacier, the classifier successfully delineates the general outline of DCG but includes a substantial number of outlying features (Figure B 11). These features are prominent in both the final classification sum image and the derived mask. As these outliers show values similar to those of the glacier tongue in the classification sum image, adjustments of the threshold could not improve the final classification mask. Notably, the DCG body of Belvedere Glacier shows two small, distinct areas of very low values in the classification sum image, resulting in two 'holes' in the final classification mask. One of these holes corresponds to a feature also present in the ground reference.

Satopanth glacier

When applying Approach III to Satopanth Glacier, the visualised results are poor, with the classifier failing to consistently capture the entire DCG area and misclassifying numerous outlying regions as DCG (Figure

B 12). The final classification sum shows medium-low values across most of the DCG area, with some higher values observed toward the southern end of the glacier. The DCG outlines can be derived fairly accurately from both the final classification sum and the mask images, as they exhibit a clear transition to even lower values surrounding large portions of the main glacier. However, as one moves away from the main glacier arms, areas with higher values follow the low-value regions, leading to misclassifications of non-DCG areas as DCG. This results in the inclusion of many outlying patches in the final classification mask. Additionally, although the overall outlines of the DCG are mapped reasonably well, there are several missing pixels within the main glacier body where the values were too low to be classified as DCG, causing false negatives.

6.3 Performance of the LST NIR Index

In a simplified classification setup, the performance of the LST NIR Index was evaluated. The LST NIR Index was used as the primary input layer for a simplified classification, supplemented by Slope and NDRI layers to provide basic additional information. The random forest (RF) classifier was configured with 50 trees and a sample size of 5'000 points. To test the added value of the combination with NIR, the LST NIR index was replaced by normalised difference indices of LST combined with different optical to shortwave infrared bands, keeping the same additional inputs (slope and NDRI) and RF structure.

When trained and tested on Zmuttgletscher for the year 2016, the LST NIR Index achieved the highest accuracy with 0.946 (Table 10). The overall accuracies were elevated due to the classifier being trained and tested on the same glacier and year. Consequently, the absolute accuracy values are not representative for broader applications. However, relative comparisons remain valid, demonstrating that the LST NIR Index outperformed other indices in this setup.

Table 10: Cross-comparison of validation accuracies of Normalised Difference Indices with LST and one additional band, and normalised LST only, with simplified RF classification.

		Validation accuracy
Normalised difference Indices	LST blue	0.941
	LST green	0.943
	LST red	0.938
	LST NIR	0.946
	LST SWIR1	0.934
	LST SWIR2	0.943
Normalised LST		0.941

The added value of combining LST with NIR was further evaluated through scatterplot visualisation, plotting LST against NIR and colouring data points by class (DCG or non-DCG) based on the SGI2016 debris cover ground reference (Figure 19). Clustering in the scatterplot showed that DCG had low to medium-high LST values but consistently low NIR values, whereas non-DCG had greater scatter across

the range. This demonstrated that the inclusion of NIR added a valuable second dimension, enhancing the ability to effectively discriminate between the two classes.

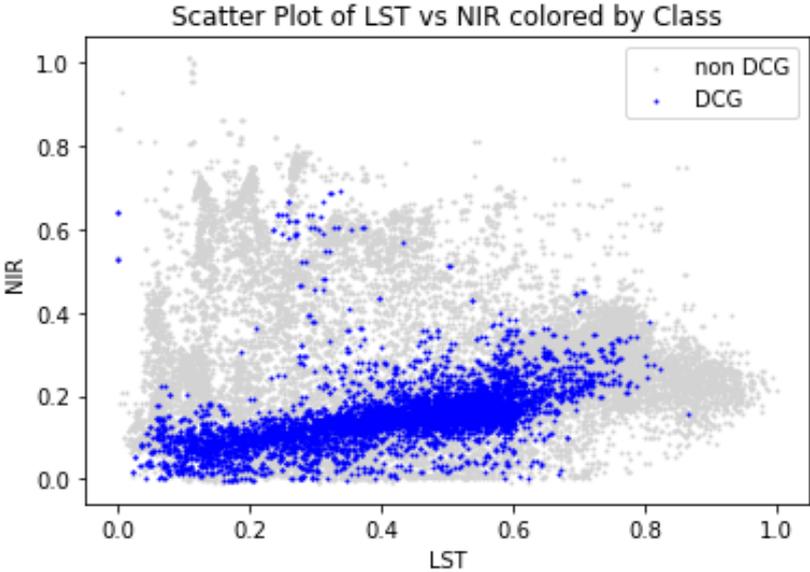


Figure 19: Scatterplot of the relation of LST and NIR for DCG and non-DCG (coloured according to SGI2016 debriscover), with data from Zmuttgletscher 2016. Visualized in Python.

7 Discussion

The discussion begins with a brief summary and interpretation of the performance of the method for the different approaches used (Section 7.1). To gain a deeper understanding of the performance of the classifier, specific results are examined through case studies (Section 7.2). The added value of the newly introduced LST-NIR index is discussed in section 7.3. Finally, section 7.4 focuses on the limitations of the methodology.

7.1 Different Approaches

The tested approaches included temporal change analysis of a single glacier over three years, and cross-glacier analysis, using the most current glacier outlines as one of the input layers (Approach I), using RGI 7.0 outlines for all years (Approach II), and excluding glacier outlines entirely from the input layers (Approach III). Overall, the results were promising in DCG delineation, with temporal change analysis and cross-glacier analysis Approach I yielding the highest accuracies. This can be explained by the greater consistency between training and testing data in the temporal analysis and the inclusion of up-to-date glacier outlines in both temporal analysis and cross-glacier analysis Approach I. Accuracies were slightly reduced in Approach II and further declined in Approach III.

As accuracy metrics refer to the correspondence between the classification results and the ground reference derived from the RGI and SGI, they must be interpreted with caution. A lower validation accuracy does not necessarily indicate poor model performance, but rather highlights a larger difference between the results and the existing glacier inventories. Determining whether these differences are due to inaccuracies in the classification results or in the ground reference requires qualitative assessments of individual cases (see Section 7.2).

7.1.1 Temporal change analysis

The methodology was initially applied to assess temporal changes in DCG extent across different years for individual glaciers. Training the classifier using data from one glacier over two years and testing it on a third benefited from the consistency in the input data, allowing evaluation under 'ideal' conditions where input variations were expected to primarily reflect changes in glacier extent.

This approach yielded high validation accuracies, ranging from 0.854 to 0.970 (Table 6), demonstrating the classifier's effectiveness in identifying DCG and accurately delineating glacier contours over time. However, an anomaly was observed in the classification of Zmuttgletscher in 2010 (Figure 14), where a large portion of the glacier was not identified by more than half of the 20 random band combinations. This anomaly underscores the limitations and uncertainties that can arise from irregularities in the input data or training process.

In all other cases analysed (including Zmuttgletscher 2003 and 2016), the classifier performed well, successfully delineating DCG outlines and identifying debris-covered areas. These results support the suitability of the methodology for temporal change detection, assuming consistent and reliable training and validation data.

This approach is constrained by the need for ground reference data for each year used in training and, if statistical evaluation is required, also for the test year. On a global scale, only the RGI dataset is available, which dates back to 2003 for most glaciers analysed in this study. Extending this temporal

change analysis approach to regions outside of Switzerland would require training the classifier using only the 2003 dataset, and not allow for statistical evaluation of the test year. This limitation halves the training dataset, increasing the influence of outliers and specific characteristics of individual years and scenes, thereby reducing the overall reliability of the classification results. Therefore, this approach was used as a first assessment of the performance of the methodology, benefiting from high consistency in training and testing datasets, but is not suitable for transfer to other regions.

7.1.2 Cross-glacier analysis

The methodology was subsequently applied across multiple glaciers to assess its robustness and potential for transferability using Approaches I, II, and III. Approaches II and III increase the transferability of the method as they are not restricted to specific years, although statistical evaluation is still only possible for years with available ground reference data.

Approach I served as an initial test of the classifier's performance when trained with data from two different glaciers instead of one, increasing the variability in the training dataset. This variation introduced a broader range of conditions, reflecting differences across glaciers in the input layers used for classification.

For instance, slope played a critical role in delineating DCG outlines for Unteraargletscher, where the flat glacier surface transitions abruptly into steep mountain flanks (Figure 20 b). In contrast, Zmuttgletscher, situated in a wider valley, although flanked by steep moraines, exhibits a slope change between the glacier surface and its surroundings that is less pronounced and does not consistently align with the glacier outlines (Figure 20 a). These differences highlight the increased complexity introduced by combining data from multiple glaciers. Testing the classifier with this approach thus provides a more comprehensive evaluation of the methodology's adaptability.

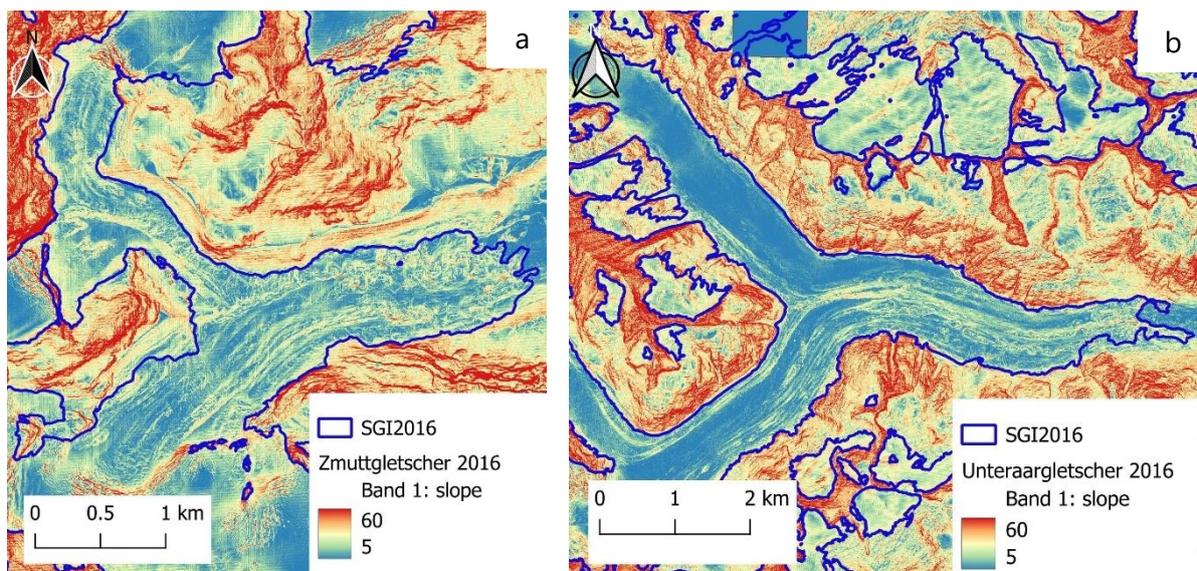


Figure 20: Visualisation of slope based on the SwissAlti3D (2019) for Zmuttgletscher (a) and Unteraargletscher (b), at a resolution of 2m, used for training and testing of classifier on the respective glaciers in 2016.

Although all three approaches used in the cross-glacier analysis performed slightly less well than the temporal change analysis, the validation results were still generally satisfactory, with most accuracies being greater than 0.8.

Approach I

For Approach I, for Zmuttgletscher, Unteraargletscher, and Oberaletschgletscher, validation accuracies ranged from 0.840 to 0.940 (Table 7).

In most cases, the final classification output effectively identified DCG and delineated glacier outlines. For Oberaletschgletscher and Zmuttgletscher, as well as Unteraargletscher in 2003, the classification results agreed well with the expected glacier shapes and ground reference data. However, in some cases the final classification sum, which aggregates the results of multiple band combinations, better reflected the true DCG outlines than the actual final classification mask produced by Otsu's thresholding method. This was particularly evident for Oberaletschgletscher, where a slightly higher threshold might have resulted in a more accurate delineation of the glacier.

The use of Otsu's threshold in this method was chosen to automate the process and avoid manual selection, which is impractical when applying the method to multiple glaciers or to years without available ground reference data. As the thresholds vary between glaciers and years (ranging from 6 to 9 for the 20 band combinations tested), a fixed threshold could not be used. In addition, the aim of the methodology was not to match the output as closely as possible to the ground reference, but rather to develop a semi-automated approach capable of refining glacier contours independently. Nevertheless, in certain cases a manually adjusted threshold might have improved the accuracy of the final classification mask.

Approach I was used as a preliminary evaluation of the classifier's performance when trained and tested on data from different glaciers, using the same input layers as in the temporal change analysis. This approach is limited to years and glaciers for which a current glacier outline is available, as the method is designed to refine these outlines rather than generate entirely new ones. Including existing outlines as input layers allows more contextual information to be incorporated into the classification process, but it also carries the risk of introducing errors from the original outlines into the refined results. While the validation accuracy may appear high due to alignment to the ground reference (which is itself based on the included outlines), the true accuracy of the refined outlines relative to the actual glacier extent remains uncertain and requires further evaluation.

Approach II

Using Approach II, the validation accuracies of the results slightly decreased using this approach, now ranging from 0.840 to 0.919 (Table 8).

In Approach II, the inclusion of the glacier inventory in the classification exploits a valuable source of data that can improve accuracy, as glacier inventories provide reliable information on glacier boundaries. As the objective of the method was to refine the boundaries, the inclusion of existing glacier inventories is well suited to this objective and shows potential for DCG delineation on a larger scale. The inclusion of the RGI 7.0 glacier outline and the lack of reliance on current glacier outlines allow for transferability of this approach to different glaciers and years but introduce some uncertainties.

When applied to a current year, the classifier must adjust the outdated 2003 glacier outline using information from other, current (year of analysis) input layers. This reliance on outdated outlines resulted in slightly decreased statistical metrics, as the ground reference and the glacier outline used as an input layer are of different origin. In addition, as most glaciers have retreated since 2003, the RGI outlines are more likely to overestimate the true extent of glaciers, as seen in the classification of Zmuttgletscher (Figure 17). This effect was not observed for Unteraargletscher or Oberaletschgletscher, suggesting that the extent of this effect may vary depending on specific glacier characteristics.

Nevertheless, the approach has significant advantages. It is highly transferable, allowing application to any glacier (as long as it is mapped in the RGI 7.0) and year, even in the absence of current outlines. Furthermore, while the inclusion of RGI outlines introduces some uncertainty, it also adds valuable prior information that improves the performance of the classifier compared to an approach without outlines.

Overall, Approach II shows promising potential for DCG delineation, although further refinement of the methodology could improve its performance. A key advantage of this approach is its transferability to other regions and years, as it incorporates some information from existing glacier outlines while remaining adaptable. However, this transferability was not explored within the scope of this thesis, as it precludes statistical evaluation due to the lack of a comparable ground reference. Instead, such applications would require a qualitative assessment of the results.

Approach III

Approach III eliminates reliance on existing glacier outlines, instead utilizing satellite imagery and topographic data. This approach avoids the potential propagation of inaccuracies from outdated inventories. However, it results in slightly less accurate delineations compared to Approach II, with validation accuracies ranging from 0.743 to 0.911 (Table 9). It is important to note that validation here refers to comparison with ground reference data derived from existing glacier inventories, which are not definitive truths but come with their own uncertainties. By not including glacier outlines, this approach is less likely to conform exactly to the extent and boundaries established in existing inventories, which naturally reduces its 'validation' accuracy in statistical terms. Nevertheless, the mapped results suggest a decrease in performance compared to previous methods. Specifically, for Zmuttgletscher, the delineated areas show irregular shapes and many patches that are likely misclassified as DCG, highlighting a notable decrease in accuracy.

Unlike the previous approaches, Approach III used 40 band combinations for the final classification, doubling the 20 combinations used previously. This adjustment took into account the reduced information available to the classifier due to the exclusion of glacier inventory outlines. Different band combinations proved to be effective depending on the input data, and increasing their number improved the robustness of the classifier in identifying DCG across glaciers and years.

However, increasing the number of band combinations doubled the computational load, exceeding the capacity of Google Earth Engine (GEE) with the original script. As a result, the threshold selection method was changed from Otsu's algorithm to manual selection based on the visualised classification sum image, simplifying the process while maintaining reasonable accuracy.

7.2 Case studies of model performance

This section provides an in-detail discussion of specific classification results or characteristics that could be observed in a number of results, to better understand the performance of the classifier and possible interpretation of its results.

7.2.1 Distinct value transitions on the main DCG body

In several cases, a distinct transition from high to low values on the main DCG tongue could be observed in the final classification sum images.

To explore potential causes for the lower values observed at the glacier tongue's front, classification sum images for Zmuttgletscher (generated with Approach I) were overlaid onto a high-resolution satellite image (Figure 21). This comparison, however, did not clarify the reason for reduced values in this area. Although the frontal glacier region does contain slightly more ice cliffs, which might influence slope and temperature data, these specific input layers did not exhibit any evident patterns explaining this transition. The classifier was trained with data from Unteraargletscher and Oberaletschgletscher and then applied to Zmuttgletscher, so this transition could stem from inconsistencies among any of the input layers of these glaciers. Yet, a review of all input layers showed no distinct outliers that could account for this effect.

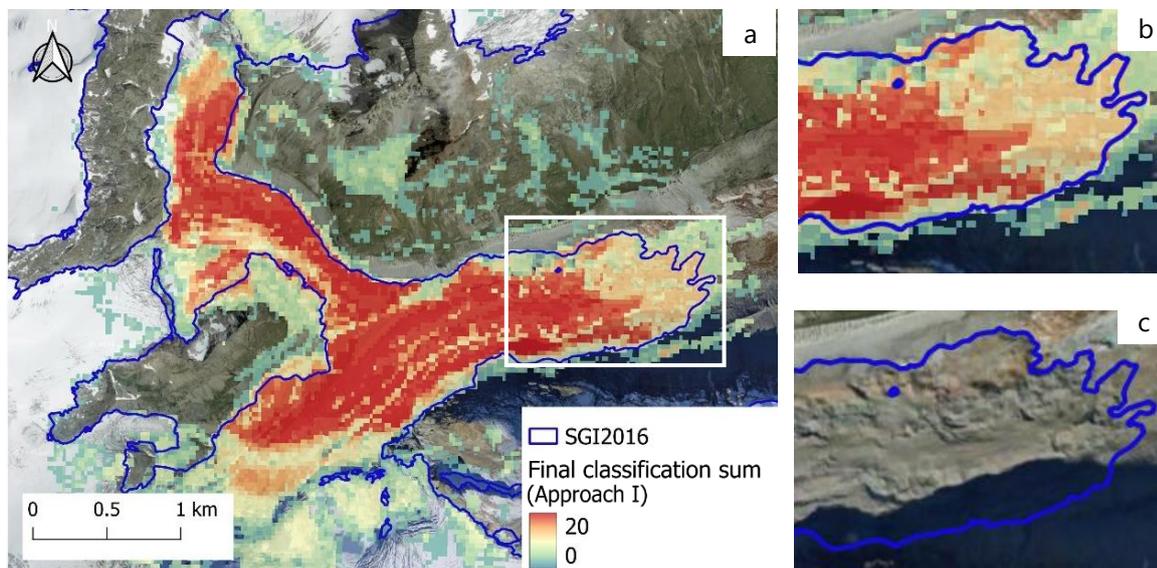


Figure 21: Classification sum of Zmuttgletscher 2016 generated with Approach I and overlaid on high-resolution google satellite imagery (a) and zoom to glacier front (b and c). Compared to the SGI2016 glacier outlines.

Similar high-to-low transitions on the main DCG body were observed in other cases. For example, Zmuttgletscher's classification sum images in 2003 and 2010, generated with Approach I (Figure 15), as well as Oberaletschgletscher's 2016 results using Approaches II (Figure B 8) and III (Figure B 10), displayed this pattern. In none of these cases did any input layer convincingly account for the observed transition.

The classification sum image of Oberaletschgletscher generated with Approach III (Figure B 10) shows a sharply defined transition, appearing as though cut off with a knife. Unfortunately, the available high-resolution satellite imagery for this area shows the glacier covered in snow, limiting the information that can be derived. However, overlaying the SGI2016 debris-cover outlines reveals a high level of detail in delineating individual ice cliffs on the glacier tongue (Figure 22).

The frontal ~1 km of the glacier tongue, which shows significantly lower values compared to the upper glacier, contains a greater density of ice cliffs. As with the previously discussed Zmuttgletscher results, these ice cliffs might influence various input layers used in the classification. Despite this, a review of all input layers did not reveal any clear explanation for the marked transition observed in the classification sum image.

This highlights the need for a more detailed analysis of the methodology's performance, including its application to a broader range of glaciers, to determine if these patterns recur and to identify their underlying causes. Despite these value transitions, the overall performance of the classifier was not significantly affected, as the thresholds for differentiating DCG from non-DCG were set low enough to capture the main DCG bodies in all cases. However, using a higher threshold could help exclude more outlying patches falsely classified as DCG, such as in the 2016 result for Oberaletschgletscher with Approach III. Identifying and, ideally, addressing the cause of these transitions would enhance the methodology by improving classification precision and reducing errors.

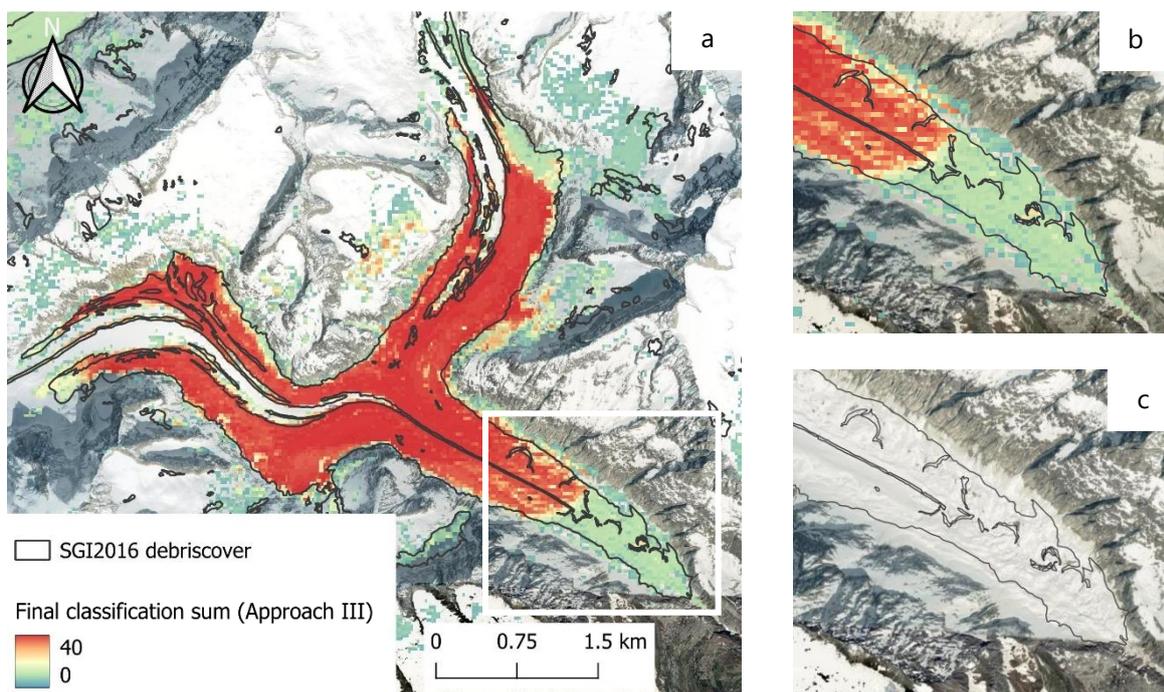


Figure 22: Classification sum of Oberaletschgletscher 2016 generated with Approach III and overlaid on high-resolution google satellite imagery (a), with zoom to glacier front (b and c). Compared to the SGI2016 debris-cover outlines.

7.2.2 Potential DCG misclassifications and 'noise'

In many cases, the final classification masks included DCG patches outside the main glacier body, which were not present in the corresponding ground references and were assumed to be misclassifications. However, some of these patches may in fact represent real DCG.

A notable example is Oberaletschgletscher 2010, which was generated based on temporal analysis and showed several distinct medium-sized outlying patches (Figure 23). Comparison of these patches with SGI2010 outlines shows that they are also delineated as glacier area in the SGI. The RGB Landsat image used as the classification baseline shows these patches primarily as snow/ice, which were excluded from the ground reference by the NDSI mask. However, the parts classified as DCG in the final mask coincide with the debris-covered parts of these patches, suggesting that the performance of the method was better than initially assumed.

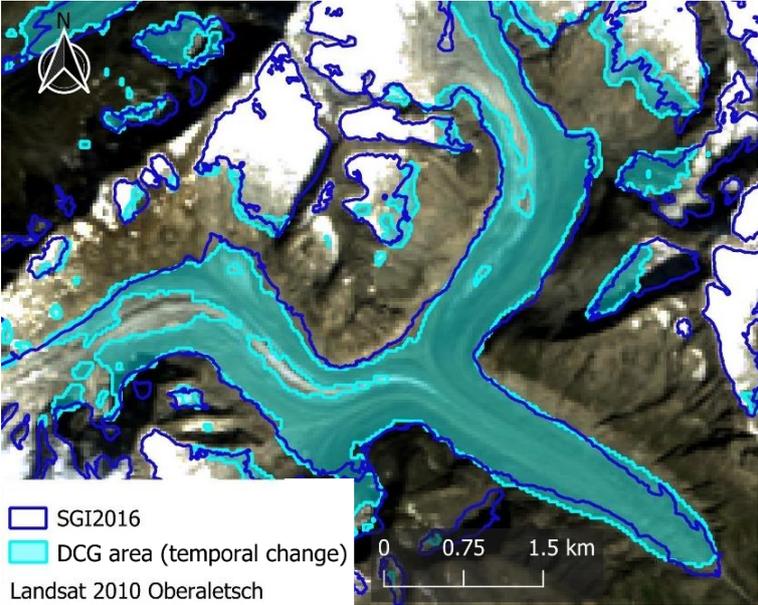


Figure 23: DCG classification result of Oberaletschgletscher 2010 overlaid on the corresponding Landsat scene and compared to SGI2010 outlines.

As the size of such patches is reduced, their correspondence with existing glacier inventories decreases, probably due to mixed pixel signals or the presence of small snow or clouds in the scene causing local variations. Single pixels or very small outlying patches were therefore treated as 'noise' or misclassifications and removed using a smoothing filter.

For Satopanth Glacier in 2003 using Approach I (Figure B 6), the final classification sum displayed numerous very small outlying patches, predominantly single pixels or small clusters. Applying a smoothing filter before vectorizing the results effectively reduced this 'noise,' enhancing the visibility of overall DCG patterns.

Upon closer inspection, this noise primarily falls on debris-covered or partially debris-covered areas outside the RGI outlines (Figure 24). Larger patches outside the RGI boundaries, which likely represent misclassifications, were not eliminated by the smoothing filter. However, removing the smaller noise significantly clarifies the results, making the output more interpretable. The refined outlines align more closely with actual DCG, reducing the influence of pixel-level variations and improving the usability of the result.

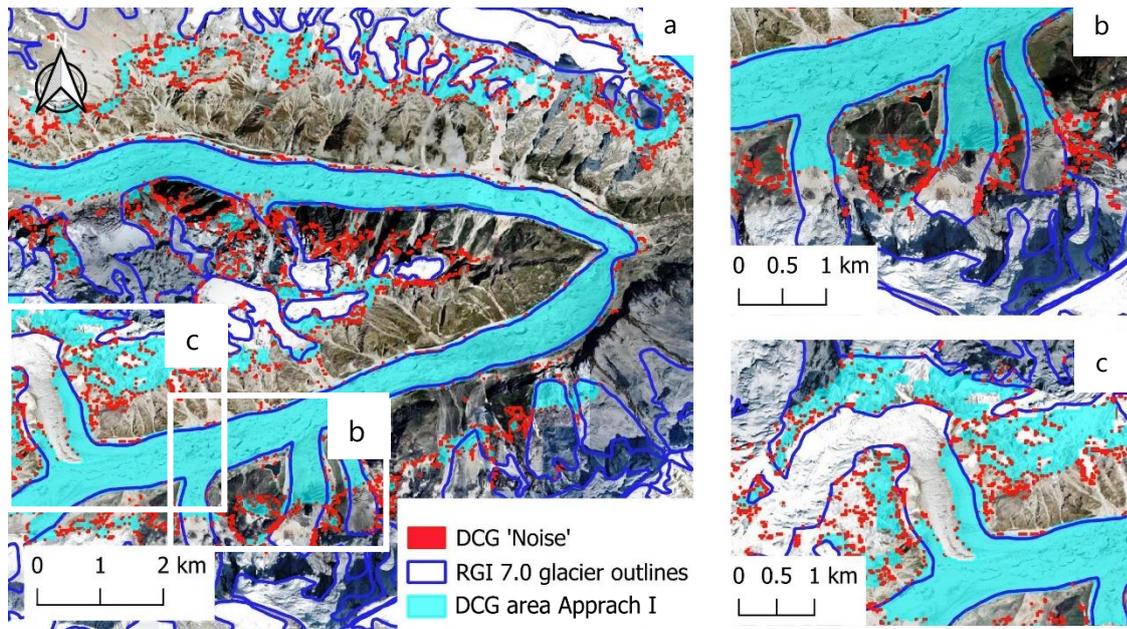


Figure 24: Classification result for Satopanth glacier with Approach I, showing the DCG outlines of the classifier compared to RGI 7.0 outlines and DCG 'noise' that was filtered out by applying a smoothing filter (a), with zoom to areas on upper south-western glacier arm (b and c).

7.2.3 Glacier terminus and wideness: the need for accurate thresholding

The actual glacier extent of the final classification result heavily depends on the threshold that was applied to the classification sum image. In many cases, there is a rather clear and distinct value transition of higher values for DCG and lower values for the surrounding. However, in some cases, distinct value transitions on the glacier, as previously discussed, or a stepwise value transition at the glacier terminus or its sides, make it less clear where the threshold to differentiate between DCG and non-DCG should be set. Using Otsu's method to determine the threshold is a way to maximise the difference between the two classes but does not consider spatial patterns. Setting the threshold manually and individually for each glacier, as was done in Approach III, does consider spatial patterns as the threshold was selected based on visualisation of the final classification sum, nevertheless it is debateable what an accurate threshold is.

The final classification sum of Approach I for Zmuttgletscher 2003 shows a stepwise transition of values at the glacier front (Figure 25 a). The threshold used for deriving the final mask image, determined using Otsu's method was 9 and included also the area of slightly lower values to be part of the DCG. Manually setting a higher threshold of e.g. 17 would exclude this area from the DCG body, marking it as non-DCG (Figure 25 b). To assess whether Otsu's threshold was set accurately, or if the drop in values could translate to a shorter glacier tongue, the results of both thresholds were tested against the ground reference and compared to high-resolution satellite imagery (Figure 26 a).

Compared to the ground reference, Otsu's threshold is more effective in capturing DCG than the increased threshold of 17 is. The accuracy of the result decreases for the higher threshold, as well as F1-score and recall (Table 11). Precision increases slightly, as less pixels are misclassified as DCG, probably addressing noise and patches outside the main DCG area.

Table 11: Comparison of statistical metrics for Zmuttgletscher 2003 with cross-glacier analysis, Approach I, when using different thresholds, either based on Otsu's method (9) or manually selected (17).

	Threshold = 9 (Otsu)	Threshold = 17
Accuracy	0.940	0.881
F1-Score	0.939	0.869
Precision	0.952	0.970
Recall	0.927	0.787

Visualizing the results derived from the two different thresholds reveals their spatial patterns and differences (Figure 25 b). The outline generated with the lower threshold more closely follows the RGI outline at the glacier front, while the higher-threshold outline aligns better with the RGI outline in the upper glacier regions, particularly at the northwestern glacier arm and along both sides of the main glacier tongue.

Comparing these outlines to high-resolution satellite imagery, the lower-threshold outline clearly exaggerates the glacier tongue's width, crossing the apparent boundary of the side moraines (Figure 25 b). Since the higher-threshold outline generally fits better overall, it raises the question of whether the glacier front might also be more accurately represented by this outline, potentially suggesting that the RGI outline depicts the front as overly extended.

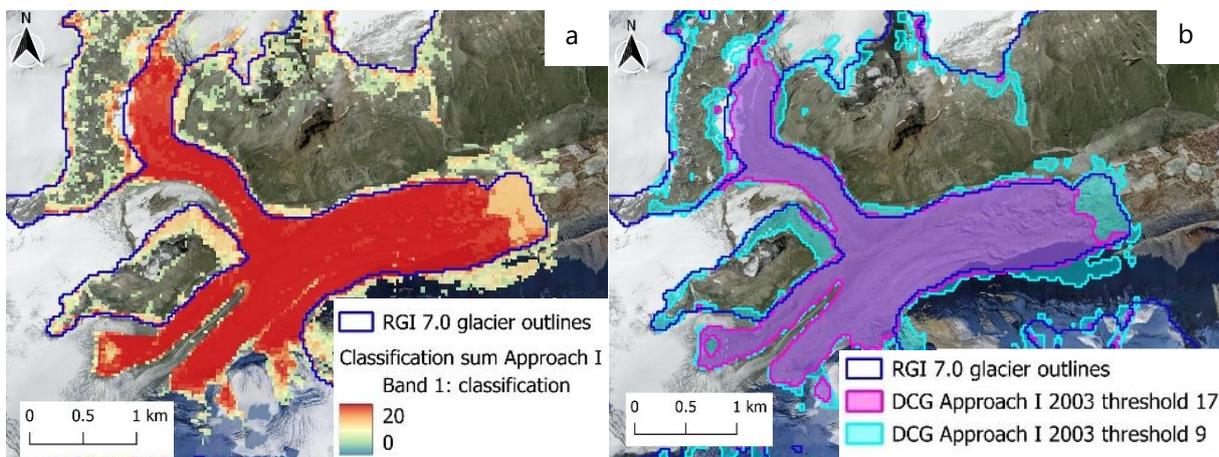


Figure 25: Final classification sum of Zmuttgletscher 2003, with Approach I (a), and derived outlines based on different thresholds (b). Overlaid on high-resolution Google satellite imagery.

To examine the patterns of outlines at the glacier front more closely, Figure 26 provides a zoom into this region. The high-resolution satellite imagery used for this comparison (Figure 26 a and b) was acquired between 26.09.2016 and 04.09.2023 (Google Earth, 2024), so it cannot be used directly to derive the extent of the glacier in 2003. What is clearly visible, however, is a distinct feature in the ground moraine, resembling a line that marks the boundary between a slightly more vegetated area in the glacier forefield and the debris-covered area closer to the glacier (Figure 26 b, b1). The fact that this transition is so distinct and clearly visible suggests that the glacier has been at this extent for some time. Whether this was the case in 2003 is more challenging to determine.

To compare the outlines with images closer to the glacier's 2003 state, they were superimposed on a 2005 aerial image (Bundesamt für Landestopografie swisstopo, 2005) as the closest available scene (Figure 26 c and d). While the distinct feature identified as b1 is less visible in this scene, four other features were identified in the region that was classified as either DCG or non-DCG depending on the threshold used (Figure 26 d, d1-d4).

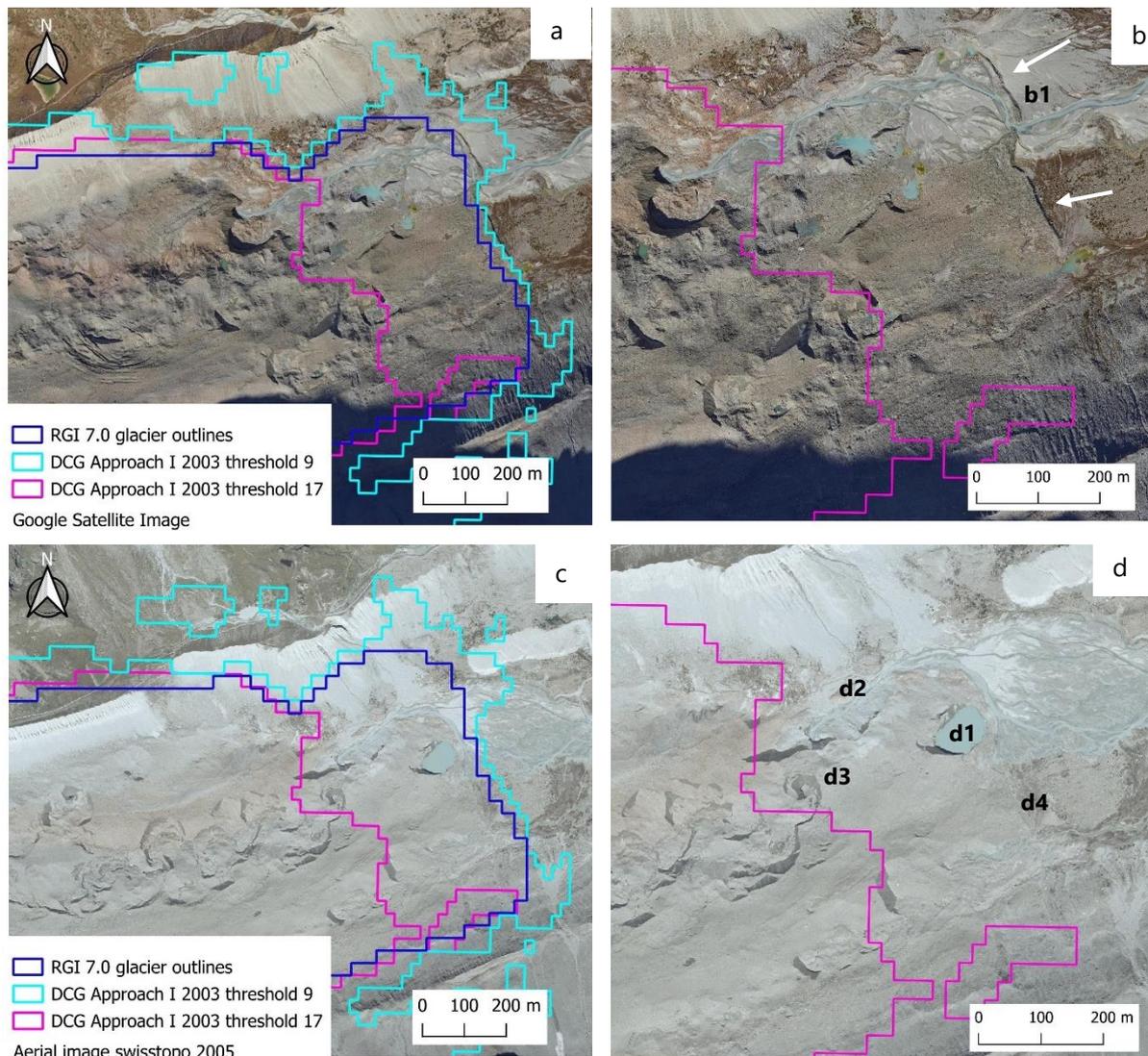


Figure 26: Different outlines for Zmuttgletscher in 2003, generated using Approach I, showing variations based on different thresholds. The outlines are overlaid on high-resolution Google satellite imagery (a and b) and a 2005 aerial image (c and d). Images b and d zoom in on the area of discussion, showing a ground moraine feature matching the RGI 7.0 terminus (b1), as well as a lake (d1), melt river (d2), possible ice cliffs (d3), and vegetation (d4).

Feature d1 appears to be a lake, which could be either a supraglacial lake or a lake in the glacier forefield. While supraglacial lakes on DCG are typically smaller than on clean ice glaciers (Zeller et al., 2022) it remains possible that this lake is supraglacial. Feature d2 shows a meltwater stream, which could also be in the glacier forefield or supraglacial. Supraglacial streams are less common on DCG than on clean-ice glaciers, and tend to be smaller in volume (Fyffe et al., 2019), but their presence on DCG cannot be ruled out.

Feature d3 suggests the presence of possible ice cliffs, which would support the classification of this area as DCG. In contrast, feature d4 appears to show vegetation, which would indicate the absence of DCG in this area. However, vegetation is not a definitive indicator, and its presence does not completely rule out DCG.

None of these features provide definitive evidence for or against the presence of DCG in this particular region. While features d1, d2 and d4 lean slightly towards a non-DCG classification and feature d3 suggests the presence of DCG, these interpretations are non-conclusive. This suggests that the true DCG outline is likely to lie somewhere between the different outlines presented. However, a precise qualitative assessment of the exact extent is beyond the scope of this paper.

7.2.4 Holes in the main DCG body: the case of Belvedere glacier

For Belvedere Glacier, holes appear in the final classification mask over the main glacier area when using both Approach I (Figure B 5) and III (Figure B 11). Examination of the Landsat scene used for classification shows that the areas classified as non-DCG correspond to pixels potentially representing snow or cloud cover. However, the coarse resolution of the scene makes it challenging to definitively identify these patches as snow, as they are only suggested by lighter-coloured pixels. In the NDSI mask, these areas were classified as snow/ice and subsequently excluded from the RGI-based ground reference. As the holes appear in both the ground reference and the classification result, this feature does not affect the validation accuracy. However, the occurrence of such non-DCG holes within the DCG body seems highly unlikely from a glaciological perspective and was therefore investigated further.

Comparison with a high-resolution satellite image, although from a more recent date, provides no evidence of permanent snow cover in these areas (Figure 27), supporting the assumption that this feature is specific to the single scene used for classification. To qualify as DCG, an area must consist of glacier ice covered by debris. The observed holes, though located near the side border of the main glacier body, are surrounded by DCG. Given the flow dynamics of glaciers, it is reasonable to expect underlying ice in these areas, as ice is present both upstream and downstream of the holes. Additionally, since the surrounding ice is debris-covered and there are no topographic barriers to glacier flow or debris transport in these regions, these areas are likely to be DCG despite their classification as non-DCG in this instance.

Comparison to the normalised LST (Figure 27 d) reveals lower temperatures in these areas, which influenced LST-based input layers, such as the LST NIR Index and its variations, leading to the observed holes in the final classification. This further supports the assumption that these features are specific to the single scene used for classification, likely resulting from localized snow cover or potentially high-altitude clouds. While such clouds are barely discernible in optical imagery, they are clearly reflected in the LST signal, potentially contributing to the classification anomalies.

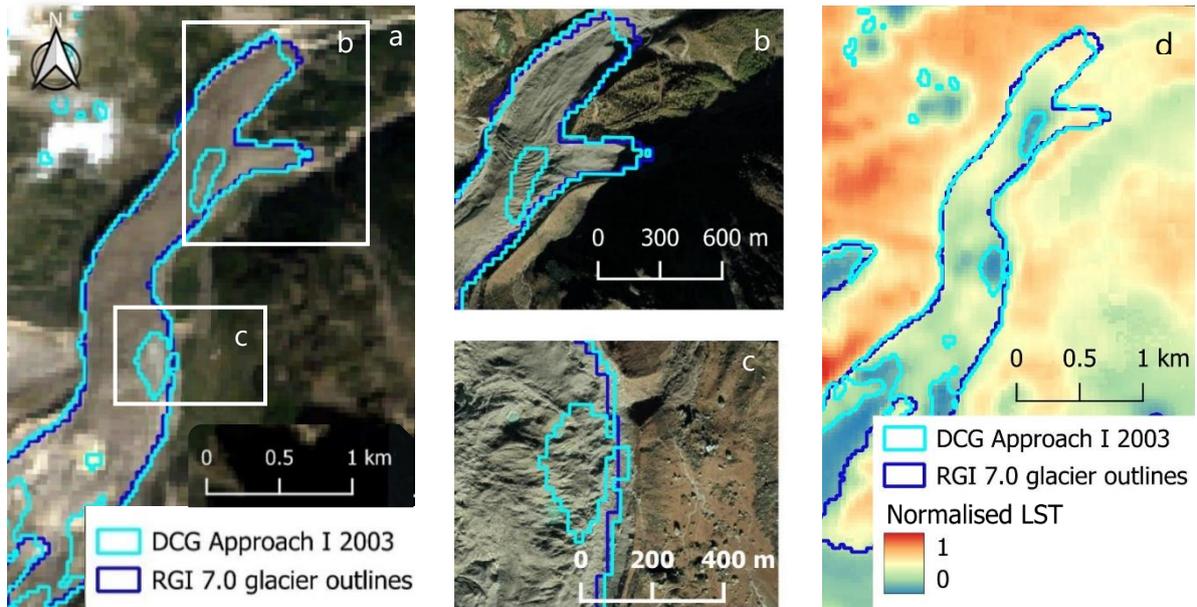


Figure 27: Classification result of Belvedere glacier 2003 (Approach I) compared to RGI 7.0 outlines, overlaid on Landsat scene used for classification (a), high-resolution Google satellite imagery (b, c), and normalised LST (d).

7.2.5 The issue with temporal change analysis on Zmuttgletscher 2010

A large hole was observed in the main DCG body of Zmuttgletscher in 2010, which was misclassified as non-DCG with the temporal change analysis (Figure 14). In this approach, the classifier was trained on data from 2003 and 2016 and applied to 2010. While this approach typically yielded high validation accuracies and strong alignment with ground reference, this case deviates from this trend. The actual absence of DCG in the region of the hole is highly unrealistic due to the dynamics of the glacier, and no such reports are found in the extensive analysis of Zmuttgletscher and its debris-cover by Mölg et al. who have documented changes in the debris-cover of the glacier over the past decades, suggesting that this result is an error in the classification.

When all input layers were examined, no significant deviations were found in neither the layers based on the Landsat scene, nor in the ones containing topographical information. The classifier performed well for Zmuttgletscher in 2010 with all other approaches tested using the same input data (Approach I uses identical input data per glacier as the temporal change analysis). This suggests that factors such as clouds or snow in the scene, which could explain problems such as those seen for Belvedere Glacier, are unlikely in this case.

One input layer that differs significantly in 2010 is the glacier outline, which according to the SGI2010 inventory is significantly narrower than the outlines in 2003 and 2016. The glacier outlines in the ground reference show a pattern of being wider in 2003, then narrower in 2010 and wider again in 2016. This progression is highly unlikely and suggests a likely discrepancy in the 2010 glacier outline (discussed further in Section 7.4.1). This discrepancy is particularly relevant because the classifier was trained on data from 2003 and 2016, where such inconsistencies were absent, and then applied to 2010. The inconsistency between the glacier outline in 2010 and other input layers such as slope, elevation and LST (none of which showed evidence of glacier narrowing) likely contributed to the poor performance of the classifier for that year, including the large area of DCG misclassified as non-DCG.

This problem was not observed with any of the other approaches, including Approach I of the cross-glacier analysis, which uses the same input layers for classification as the temporal change analysis. However, in the cross-glacier analysis, the training data for classification came from Unteraargletscher and Oberaargletscher in 2010, where the glacier outlines were also slightly narrower than in other years for the respective glaciers (as the ground reference for 2010 was consistently based on SGI2010). This alignment of the glacier outline standards reduced inconsistencies between the training and test datasets in the cross-glacier analysis, probably leading to the observed improved performance of the classifier for Zmuttgletscher in 2010 with this approach.

This discrepancy between the glacier outlines and other data may have contributed to the large area of DCG being misclassified as non-DCG. However, the specific placement of the misclassified hole in the middle of the glacier cannot be fully explained by this inconsistency, leaving some uncertainty as to the exact cause.

7.3 Benefit of LST NIR Index

To improve the detection of DCG, this method used a normalised difference index combining the LST and NIR bands. This index allowed improved differentiation of DCG from surrounding surfaces such as snow, ice, soil, vegetation and rocky terrain by utilising the reflective qualities and higher spatial resolution of the NIR band alongside the thermal signal of the LST band.

Incorporating NIR into the LST-NIR index offers several advantages for DCG mapping:

- Differentiation: NIR helps to distinguish DCG from surrounding features such as snow/ice, soil, vegetation and, to some extent, rocky surfaces, improving the accuracy of the classification.
- Second dimension: The combination of LST and NIR introduces a second dimension to the dataset, helping to better differentiate between DCG and non-DCG by exploiting the different spectral and thermal characteristics of surfaces.
- Validation accuracy: A cross-analysis tested the performance of different band combinations with normalised LST using a random forest classifier, with all other input variables held constant. Combining LST with NIR achieved the highest validation accuracy, and provided most consistent and complete data when mapped, ensuring reliable classification under different conditions. Combining LST with any additional band enhanced the visual differentiation between DCG and non-DCG by providing a second dimension of information. This improvement was largely due to the higher spatial resolution of the additional bands, which at 30 m offer finer detail compared to the coarser resolution of LST. Among the tested combinations, which calculated the normalised difference between LST and blue, red, green, NIR, SWIR 1, and SWIR 2 bands, respectively, the pairing of LST with the NIR band demonstrated the greatest improvement in distinguishing DCG from non-DCG, highlighting the value of the NIR band for this purpose.

7.4 Limitations

7.4.1 Coarse resolution and mixed pixels

While a resolution of 30 m is considered high resolution in satellite remote sensing, it is relatively coarse for capturing the small-scale changes required to delineate DCG outlines in alpine glacier environments. The LST data, originally at 100 m resolution and resampled to 30 m, face inherent limitations in

representing fine detail. Each 30 m pixel contains mixed reflectance signals from different surface characteristics, and at 100 m this issue is exacerbated, resulting in greater loss of information.

Glacier outlines, where distinct changes in the LST signal are expected, are particularly susceptible to these mixed pixel effects. Consequently, resolution limitations prevent the developed methodology from achieving the level of detail and precision found in datasets such as the SGI2016.

Figure 28 illustrates the challenges posed by 30 m and 100 m resolution data in accurately delineating DCG outlines. Subfigure (b) shows an RGB image of the Landsat scene used for classification, while (a) shows a high-resolution Google satellite image of the glacier front for comparison. Features such as ice cliffs and moraine structures, which are sharply defined in the high-resolution imagery, become blurred and blended into mixed pixels at 30 m resolution, demonstrating the limitations in capturing the fine scale detail required for accurate DCG mapping. The visualised LST of the same area (c), originally at 100 m resolution but resampled to 30 m, shows even less information on where to draw DCG outlines.

The challenges of using open source thermal data for DCG mapping due to its coarse resolution have also been noted in other studies (Jawak et al., 2022; Karimi et al., 2012), and although partially addressed in this thesis by integrating multiple layers of information, this remains a limiting factor.

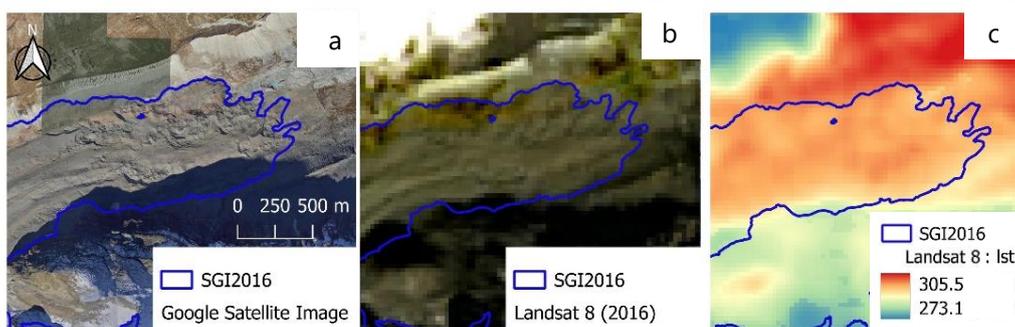


Figure 28: Exemplary visualisation of the glacier tongue of Zmuttgletscher 2016 with (a) high-resolution Google satellite imagery, (b) Landsat 8 RGB of 30 m resolution, and (c) Landsat 8 LST [$^{\circ}$ K] of 100m resolution resampled to 30m, illustrating the challenges in DCG delineation with data of coarse resolution.

7.4.1 Ground reference

An important limitation of this study is the accuracy of the ground reference data. The RGI 7.0, SGI2010 and SGI2016 glacier inventories are considered the true glacier outlines against which the performance of the method is assessed through visual comparison and statistical evaluation. However, these glacier outlines have inherent uncertainties.

In the case of Zmuttgletscher, the ground reference inventories indicate that the glacier area decreased between 2003 and 2010, but then increased again by 2016 (Figure 29). It is unlikely that the glacier would narrow and then widen in such a short period of time, suggesting potential inaccuracies in the glacier outlines used as ground reference. In a comprehensive analysis of the evolution of Zmuttgletscher over the past decades by Mölg *et al.* (2019), no narrowing of the glacier tongue in 2010 is indicated.

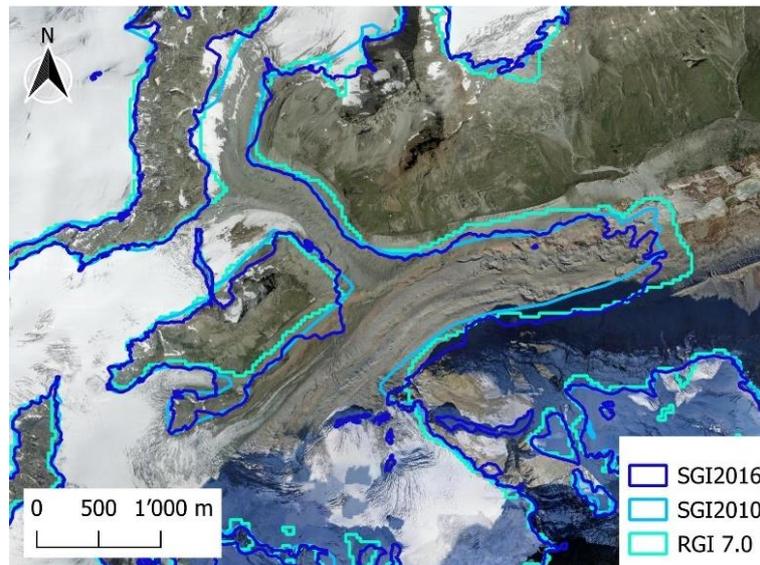


Figure 29: Comparison of the Glacier Inventory outlines for Zmuttgletscher, showing whole glacier outlines for the years 2003 (RGI 7.0), 2010 (SGI2010) and 2016 (SGI2016). A narrowing of the glacier tongue can be observed for the outline of 2010.

In 2003 and 2016, the glacier inventories did not include a specific debris-cover class, requiring the glacier outlines to be masked with an NDSI mask to derive the DCG area. This process introduced an additional layer of uncertainty into the ground reference, as the NDSI mask itself has an inherent uncertainty component, increasing the overall uncertainty of the resulting DCG delineation.

As a result, the statistical evaluation, while useful as an initial quantitative assessment of model performance, must be interpreted with caution. It remains unclear whether lower validation accuracies are due to uncertainties in the results of the methodology or uncertainties in the existing inventories used as a ground reference. Nevertheless, the SGI2016, which benefits from extensive expert input and high-resolution imagery for DCG delineation (Linsbauer et al., 2021), is expected to be more accurate than the results of this thesis, which rely on a more automated approach and lower resolution data. In contrast, the SGI2010, delineated by a single expert (Fischer et al., 2014), is likely to be less objective and may contain greater uncertainties than the SGI2016. Similarly, the RGI 7.0, which prioritises global coverage over precision at the level of individual glaciers, has uncertainties that are particularly pronounced for the DCG (RGI 7.0 Consortium, 2023).

7.4.2 Debris thickness

Another factor contributing to uncertainties in the method is the debris thickness of SGD. As discussed in section 2.3 (Influence of debris thickness on LST), debris layers greater than 50 cm can attenuate the thermal signal from the underlying ice, reducing the effectiveness of LST in distinguishing SGD from PGD (Bolch et al., 2007; Ranzi et al., 2004). Precise data on debris thickness distribution are not available for the glaciers analysed. However, it is known that debris thickness typically increases down-glacier as ice flow velocity decreases.

The results were generally more accurate in delineating DCG up-glacier and less accurate down-glacier towards the terminus. However, this can mainly be explained by the characteristics of non-DCG in these areas. While in higher areas the clear transition between snow/ice and debris makes classification easier,

towards the glacier terminus the distinction between DCG and non-DCG becomes more difficult due to the debris covering both the glacier and its forefield. This highlights the challenge that the methodology was designed to address. Thus, the influence of debris thickness on the classification results remains unclear.

However, the inclusion of additional input layers in the classification process is assumed to improve the robustness of the method to local variations in debris thickness. No specific areas of the analysed glaciers were consistently misclassified across the different approaches, suggesting that variations in debris thickness do not significantly affect the results. For example, although Mölg *et al.* (2019) reported higher debris thickness on the elongated southern ridge of Zmuttgletscher, no anomalies were observed in the classification results for this region.

For Unteraargletscher, a debris thickness of 10–20 cm was reported (GLAMOS, 2022; Huss *et al.*, 2007). While this does not enable precise evaluation of model performance relative to debris thickness, the method performed well overall in the case of Unteraargletscher. It remains unclear to what extent this success is influenced by the relatively thin debris cover.

More detailed information on the distribution of debris thickness on DCGs would be valuable for assessing the contribution of debris thickness to uncertainties in the results. However, in the absence of detailed debris thickness data for the glaciers analysed, the exact impact of this factor on the results of the method remains unclear.

7.4.3 Other influences on LST

Several factors other than the underlying ice influence LST, which can affect its reliability in DCG delineation. Elevation plays an important role, as LST tends to decrease with increasing altitude (Malbêteau *et al.*, 2017). This effect has been addressed by normalising LST values using a regression-based elevation correction. Slope and aspect also influence LST, with steeper slopes and south-facing aspects typically showing higher LST values. These variables were included in the classification process to account for their effects, following the approach of previous studies that used thermal data to discriminate between SGD and PGD while also incorporating topographic information. Mitkari *et al.* (2022) and Karimi *et al.* (2012) included slope in their analyses but used region-specific thresholds. A key advantage of the methodology presented in this thesis is that it eliminates the need for manual threshold adjustments for each region, making the process more automated and scalable.

Other factors include shadows, snow and cloud cover, which can distort LST and other optical reflectance values. These problems were mitigated by carefully inspecting individual scenes and selecting those free of such disturbances. However, daily variations in LST caused by changing weather conditions remain a source of uncertainty. While the selection of scenes from summer months minimises seasonal variability, this does not fully address daily variability. A potential solution could be to create an averaged LST image from several cloud- and snow-free scenes. Unfortunately, the availability of such images is limited, with only one suitable scene per summer representing the entire glacier without snow, cloud, or shadow interference in most regions used in this analysis. Kaushik *et al.* (2022) observed misclassifications where shadows were incorrectly identified as DCG. Other studies (Karimi *et al.*, 2012) have addressed this issue by distinguishing between shadowed and illuminated areas and performing separate classifications of DCG and non-DCG for each. Incorporating a similar approach into this methodology may provide an opportunity to improve the accuracy of DCG delineation, particularly in regions where shadows significantly affect image availability and classification results.

In alpine environments, several additional factors can influence LST and limit the precision of DCG mapping. Vegetation in the glacier forefield or surrounding areas alters surface temperatures, particularly during the summer months, as dense vegetation absorbs and retains heat differently than bare rock or ice (Holzman et al., 2021; Li et al., 2013). This variation, which could mislead DCG mapping, was addressed by incorporating the NDVI index into the classification. Surface moisture also plays a role (Bechtel, 2015; Holzman et al., 2021); wet or saturated surfaces, such as meltwater ponds or wet debris, tend to exhibit lower LST values compared to dry debris. This variability, particularly near glacier termini or in areas of significant melting, was mitigated by including the NDWI index. Additionally, albedo variations across different rock types influence LST, as lower albedo surfaces absorb more solar radiation, resulting in higher temperatures (Azzoni, 2017). Such local and inter-glacier variations in albedo can further complicate LST interpretation.

These factors underscore the importance of carefully considering the complex and dynamic surface processes that influence thermal signals, even though LST remains a valuable indicator for DCG mapping.

8 Conclusion and Outlook

8.1 Conclusion

The aim of this thesis was to develop a methodology for refining DCG outlines by exploiting the different thermal signals between SGD and PGD. This methodology was applied to analyse temporal changes in DCG across three years and extended to delineate outlines for different glaciers in a cross-glacier analysis.

The thesis addressed the following research questions:

RQ1 How can land surface temperature data be used effectively to distinguish between periglacial and supraglacial debris in glaciated regions?

The methodology developed incorporates LST to distinguish between PGD and SGD for DCGs. While LST alone is insufficient to make this distinction due to its coarse resolution, resulting in mixed pixels, and the influence of factors such as surface temperature variability and debris thickness, it proved valuable when combined with additional input data. This integration allowed LST to contribute effectively to the mapping of DCG extent and the delineation of outlines.

The methodology achieved generally satisfactory results, with overall high validation accuracies in delineating DCG areas. However, some limitations remain. The classifier occasionally misclassified outlying patches or areas in the glacier forefield as DCG or failed to identify some DCG regions on the main glacier body. The causes of these misclassifications could not always be determined, highlighting the need for further refinement of the method to improve its robustness and reliability.

RQ2 What is the potential for applying the developed methodology to refine glacier outlines and assess glacier changes of debris-covered glaciers across different regions, and what are its limitations in terms of scalability and robustness?

The methodology was applied to five different glaciers to assess its potential in different regions and time periods. Only open-source data was used to ensure global applicability to DCGs. Two approaches, Approach II and Approach III, were developed and tested for their transferability to map DCG in different regions. Of these, Approach II consistently outperformed Approach III in both validation accuracy and quality of visualised results across all glaciers and years tested. Approach II is therefore recommended for future applications. The developed methodology, specifically Approach II, shows promise as an additional tool for refining glacier outlines for any year and glacier for which data are available. By using open-source data, the method can be applied globally and across different time periods.

Although the developed method has proven its usefulness, there are still several limitations that may affect its scalability and robustness. The methodology still relies on visual inspection of results, and manual adjustment of thresholds could in some cases improve the effective discrimination between DCG and non-DCG.

A significant challenge is the reliance on coarse resolution LST data, which often results in mixed pixels that obscure the distinction of DCG contours. In addition, variations in debris thickness can attenuate the thermal signal from the underlying ice, making it difficult to interpret LST data to refine DCG contours. Topographic effects such as slope and aspect also influence LST, further limiting its direct applicability without complementary data sources.

To address these issues, the methodology incorporates a variety of additional input data to improve classification accuracy and mitigate the limitations of LST. The use of multiple band combinations and their integration into a composite classification framework has been shown to improve the robustness of the methodology across different regions. This approach accounts for the variability in DCG characteristics between different glacier types and climatic settings.

Validation of the methodology was limited by the availability of ground reference data, which restricted assessments to specific glaciers and time periods. Extending the validation dataset to a wider range of glacier types and climatic conditions will be crucial to further improve the scalability and robustness of the approach.

RQ4 How can the developed methodology complement existing glacier inventories, such as the Randolph Glacier Inventory (RGI) or the Swiss Glacier Inventory (SGI), to improve the accuracy of DCG delineation?

The developed methodology addresses a key limitation of the Randolph Glacier Inventory (RGI), which lacks explicit outlines of DCGs, by providing a means to delineate these areas. This extension could refine RGI data for both historical periods (~2000, consistent with RGI outlines) and more recent years, providing a valuable update to its global dataset.

In contrast, the Swiss Glacier Inventory (SGI) achieves high precision in glacier mapping through 3D visualisation and expert analysis but is limited to Switzerland. While the developed method cannot match the detail of the SGI due to its reliance on open-source satellite imagery, it provides a scalable, automated alternative suitable for global application. It can serve as an efficient means of delineating DCG, particularly in regions where extensive expert work or high-resolution datasets are not available.

8.2 Outlook

The method will benefit from forthcoming advances in remote sensing technologies, in particular the expected launch of new satellite missions by 2028. These missions will provide higher resolution TIR data, both in spatial detail and with more frequent revisits, which is particularly important in high alpine environments where cloud cover often obscures observations.

An important next step is to extend the methodology, in particular Approach II, to analyse more recent years and a wider range of glaciers. This would allow further refinement and validation of its effectiveness across different glacier types and regions. However, the current limitations of ground reference data, particularly on a global scale, limit the ability to perform quantitative statistical analyses on contemporary datasets. To address this, the methodology would benefit from targeted qualitative assessments of model performance for selected glaciers in recent years, providing insights into its adaptability and reliability under varying conditions.

Currently, the methodology relies on individual Landsat scenes as the data base, which have been manually selected to ensure that they are free of clouds, snow, shadows and other disturbances. A potential improvement could be the implementation of a composite-based approach, which would increase automation and potentially improve the availability of suitable input data. This approach could address situations where a glacier is partially clouded in several scenes, allowing the construction of a complete dataset by combining cloud-free sections from different scenes.

In addition, the use of LST values averaged over an entire summer season could be explored to mitigate the daily variability caused by changing weather conditions. While this approach is currently limited by the availability of suitable LST data, upcoming satellite missions are expected to improve not only the spatial but also the temporal resolution of LST datasets. These advances would allow the generation of more robust seasonal composites, further improving the accuracy and reliability of the methodology.

These refinements and extensions would not only refine the methodology, but also position it as a valuable tool for integrating new remote sensing datasets into glacier monitoring frameworks. By filling gaps in existing glacier inventories and addressing current limitations in DCG delineation, the methodology can contribute to a more accurate and comprehensive understanding of glacier dynamics at both regional and global scales.

Acknowledgements

I would like to thank my supervisors from the Remote Sensing Laboratories at the University of Zürich for their invaluable guidance and support throughout this thesis. Firstly, I would like to thank **Gabriele Bramati** for his patience in addressing my many questions and for guiding me through the entire process. I also wish to thank **Dr. Kathrin Naegeli** for her expertise and insightful feedback, which greatly contributed to the refinement of this work. Special thanks go to **Dr. Hendrik Wulf** for his assistance with Google Earth Engine and his creative problem-solving ideas. Their combined support has been instrumental in the completion of this thesis.

I also thank **Dr. Andreas Linsbauer** from the Glaciology and Geomorphodynamics group at the University of Zürich for his valuable insights into the SGI products.

References

- Acharya, T. D., & Yang, I. (2015). Exploring Landsat 8. *International Journal of IT, Engineering and Applied Sciences Research (IJIEASR)*, 4(4), 4–10.
- Agrawala, S., Raksakulthai, V., van Aalst, M., Larsen, P., Smith, J., & Reynolds, J. (2003). Development and Climate Change in Nepal: Focus on Water Resources and Hydropower. *Organisation for Economic Co-Operation and Development (OECD)*.
- Aguayo, R., Maussion, F., Schuster, L., Schaefer, M., Caro, A., Schmitt, P., Mackay, J., Ultee, L., Leon-Muñoz, J., & Aguayo, M. (2023). *Assessing the glacier projection uncertainties in the Patagonian Andes (40–56° S) from a catchment perspective*. *Glaciers/Glacier Hydrology*.
<https://doi.org/10.5194/egusphere-2023-2325>
- Alexander, C. (2020). Normalised difference spectral indices and urban land cover as indicators of land surface temperature (LST). *International Journal of Applied Earth Observation and Geoinformation*, 86, 102013. <https://doi.org/10.1016/j.jag.2019.102013>
- Alifu, H., Tateishi, R., & Johnson, B. (2015). A new band ratio technique for mapping debris-covered glaciers using Landsat imagery and a digital elevation model. *International Journal of Remote Sensing*, 36(8), 2063–2075.
- Attaullah, H., Khan, A., Khan, M., Atta, H., & Iqbal, M. S. (2023). Implications of Accuracy of Global Glacier Inventories in Hydrological Modeling: A Case Study of the Western Himalayan Mountain Range. *Water*, 15(22), 3887. <https://doi.org/10.3390/w15223887>
- Azzoni, R. S. (2017). *The dark side of the ice: Glaciological and biological aspects of supraglacial debris* [Doctoral Thesis, Università degli Studi di Milano]. https://doi.org/10.13130/r-s-azzoni_phd2017-02-24
- Barsi, J. A., Schott, J. R., Hook, S. J., Raqueno, N. G., Markham, B. L., & Radocinski, R. G. (2014). Landsat-8 Thermal Infrared Sensor (TIRS) Vicarious Radiometric Calibration. *Remote Sensing*, 6(11), 11607–11626. <https://doi.org/10.3390/rs61111607>
- Bauder, A., Funk, M., & Gudmundsson, G. H. (2003). The ice-thickness distribution of Unteraargletscher, Switzerland. *Annals of Glaciology*, 37, 331–336. <https://doi.org/10.3189/172756403781815852>
- Bechtel, B. (2015). A New Global Climatology of Annual Land Surface Temperature. *Remote Sensing*, 7, 2851–2870. <https://doi.org/10.3390/rs70302850>
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Biddle, D. (2015). *Mapping debris-covered glaciers in the Cordillera Blanca, Peru: An object-based image analysis approach*. [University of Louisville]. <https://ir.library.louisville.edu/etd/2220>
- Bolch, T., Buchroithner, M., Kunert, A., & Kamp, U. (2007). Automated delineation of debris-covered glaciers based on ASTER data. In *Geoinformation in Europe, Proceedings 27th Annual Symposium European Association of Remote Sensing Laboratories (EARSeL)* (pp. 403–410). Millpress.
- Buckley, S. M., Agram, P. S., Belz, J. E., Crippen, R. E., Gurrola, E. M., Hensley, S., Kobrick, M., Lavalley, M., Martin, J. M., Neumann, M., Nguyen, Q. D., Rosen, P. A., Shimada, J. G., Simard, M., & Tung, W. W. (2020). *NASADEM*. National Aeronautics and Space Administration, Jet Propulsion Laboratory.

- Bundesamt für Landestopografie swisstopo. (2005). *Aerial Images swisstopo color* (Version 20050817) [Dataset].
https://map.geo.admin.ch/#/map?lang=en¢er=2616828.2,1094051&z=8&topic=ech&layers=ch.swisstopo.lubis-luftbilder_farbe@year=all@features=lubis-luftbilder_farbe_000-293-635&bgLayer=ch.swisstopo.pixelkarte-farbe&featureInfo=default
- Bundesamt für Landestopografie swisstopo. (2022). *swissALTI3D Das hoch aufgelöste Terrainmodell der Schweiz*. Schweizerische Eidgenossenschaft.
- CGI-CNR (Comitato Glaciologico Italiano & Consiglio Nazionale delle Ricerche). (2024). *Inventory of Italian Glaciers*. CGI Italian Glaciers BETA Version V.1.
https://repo2.igg.cnr.it/ghiacciaiCGI/ghiacciai_new.html
- Colombero, C., Comina, C., De Toma, E., Franco, D., & Godio, A. (2019). Ice Thickness Estimation from Geophysical Investigations on the Terminal Lobes of Belvedere Glacier (NW Italian Alps). *Remote Sensing*, 11(7), 805. <https://doi.org/10.3390/rs11070805>
- Copernicus. (2018). *Copernicus DEM* [Dataset]. Deutsches Zentrum für Luft- und Raumfahrt (DLR) e.V. 2010-2014 and Airbus Defence and Space GmbH 2014-2018 provided under COPERNICUS by the European Union and ESA.
- Dash, P., Göttsche, F.-M., Olsesen, F. S., & Fischer, H. (2002). Land Surface Temperature and Emissivity Estimation from Passive Sensor Data: Theory and Practice-Current Trends. *International Journal of Remote Sensing*, 23(13), 2563–2581. <https://doi.org/10.1080/01431160110115041>
- De Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., & Crombé, P. (2013). Application of the topographic position index to heterogeneous landscapes. *Geomorphology*, 186, 39–49.
<https://doi.org/10.1016/j.geomorph.2012.12.015>
- Earth Resources Observation and Science (EROS) Center. (2020). *Collection-2 Landsat 8-9 OLI (Operational Land Imager) and TIRS (Thermal Infrared Sensor) Level-1 Data Products* [Other]. U.S. Geological Survey. <https://doi.org/10.5066/P975CC9B>
- ESA. (2020, November 16). *LSTM (Land Surface Temperature Monitoring) Copernicus*. eoPortal. <https://www.eoportal.org/satellite-missions/lstm#references>
- ESA. (2022, December 2). *TRISHNA (Thermal infraRed Imaging Satellite for High-resolution Natural resource Assessment)*. eoPortal. <https://www.eoportal.org/satellite-missions/trishna>
- Fahrland, E., Paschko, H., Jacob, P., & Kahabka, H. (2022). *Copernicus Digital Elevation Model, Product Handbook*. Airbus Defence and Space GmbH.
- Fischer, M., Huss, M., Barboux, C., & Hoelzle, M. (2014). The new Swiss Glacier Inventory SGI2010: Relevance of using high-resolution source data in areas dominated by very small glaciers. *Arctic, Antarctic, and Alpine Research*, 46(4), 933–945.
- Fyffe, C. L., Brock, B. W., Kirkbride, M. P., Mair, D. W. F., Arnold, N. S., Smiraglia, C., Diolaiuti, G., & Diotri, F. (2019). *Do debris-covered glaciers demonstrate distinctive hydrological behaviour compared to clean glaciers?* <https://doi.org/10.17863/CAM.36254>
- GEE. (2023a). *Convolutions | Google Earth Engine* [Based on work created and shared by Google and used according to terms described in the Creative Commons 4.0 Attribution License]. Google for Developers. https://developers.google.com/earth-engine/guides/image_convolutions

- GEE. (2023b). *Ee.Image.gradient* | *Google Earth Engine* [Based on work created and shared by Google and used according to terms described in the Creative Commons 4.0 Attribution License]. Google for Developers. <https://developers.google.com/earth-engine/apidocs/ee-image-gradient>
- GEE. (2023c). *Ee.Image.normalizedDifference* | *Google Earth Engine* [Based on work created and shared by Google and used according to terms described in the Creative Commons 4.0 Attribution License]. Google for Developers. <https://developers.google.com/earth-engine/apidocs/ee-image-normalizeddifference>
- GEE. (2023d). *Ee.Terrain.aspect* | *Google Earth Engine* [Based on work created and shared by Google and used according to terms described in the Creative Commons 4.0 Attribution License]. Google for Developers. <https://developers.google.com/earth-engine/apidocs/ee-terrain-aspect>
- GEE. (2023e). *Ee.Terrain.slope* | *Google Earth Engine* [Based on work created and shared by Google and used according to terms described in the Creative Commons 4.0 Attribution License]. Google for Developers. <https://developers.google.com/earth-engine/apidocs/ee-terrain-slope>
- GEE. (2024). *Google for Developers* [Based on work created and shared by Google and used according to terms described in the Creative Commons 4.0 Attribution License]. Google Earth Engine (GEE) Documentation. <https://developers.google.com/earth-engine/apidocs>
- GLAMOS. (2022). *The Swiss Glaciers 1880-2021/22* [Glaciological Reports No 1-142, Yearbooks of the Cryospheric Commission of the Swiss Academy of Sciences (SCNAT), published since 1964]. VAW / ETH Zurich. https://doi.glamos.ch/pubs/glrep/glrep_series.html
- GLAMOS. (2023). *The Swiss Glaciers 1880-2022/23* [Glaciological Reports No 1-142, Yearbooks of the Cryospheric Commission of the Swiss Academy of Sciences (SCNAT), published since 1964]. VAW / ETH Zurich. doi:10.18752/glrep_series
- Hall, D. K., Riggs, G. A., & Salomonson, V. V. (1995). Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sensing of Environment*, 54(2), 127–140. [https://doi.org/10.1016/0034-4257\(95\)00137-P](https://doi.org/10.1016/0034-4257(95)00137-P)
- Haranadhdy, B. R., & Karthikeyan, P. R. (2022). Classification of Fire and Smoke Images using Decision Tree Algorithm in Comparison with Logistic Regression to Measure Accuracy, Precision, Recall, F-score. *2022 14th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*, 1–5. <https://doi.org/10.1109/MACS56771.2022.10022449>
- Herreid, S., & Pellicciotti, F. (2020). The state of rock debris covering Earth's glaciers. *Nature Geoscience*, 13(9), 621–627. <https://doi.org/10.1038/s41561-020-0615-0>
- Holzman, M. E., Rivas, R. E., & Bayala, M. I. (2021). Relationship between TIR and NIR-SWIR as Indicator of Vegetation Water Availability. *Remote Sensing*, 13(17), 3371. <https://doi.org/10.3390/rs13173371>
- Huang, Q., & Cai, Y. (2009). Mapping Karst Rock in Southwest China. *Mountain Research and Development*, 29(1), 14–20. <https://doi.org/10.1659/mrd.857>
- Hulley, G., & Hook, S. (2015). *The ASTER Global Emissivity Database (ASTER GED)* (No. Version 3.0). Jet Propulsion Laboratory, California Institute of Technology.
- Huss, M., Sugiyama, S., Bauder, A., & Funk, M. (2007). Retreat Scenarios of Unteraargletscher, Switzerland, Using a Combined Ice-Flow Mass-Balance Model. *Arctic, Antarctic, and Alpine Research*, 39(3), 422–431. [https://doi.org/10.1657/1523-0430\(06-036\)\[HUSS\]2.0.CO;2](https://doi.org/10.1657/1523-0430(06-036)[HUSS]2.0.CO;2)

- Jawak, S. D., Wankhede, S. F., Luis, A. J., & Balakrishna, K. (2022). Multispectral Characteristics of Glacier Surface Facies (Chandra-Bhaga Basin, Himalaya, and Ny-Ålesund, Svalbard) through Investigations of Pixel and Object-Based Mapping Using Variable Processing Routines. *Remote Sensing*, *14*(24), 6311. <https://doi.org/10.3390/rs14246311>
- Jimenez-Munoz, J. C., Cristobal, J., Sobrino, J. A., Soria, G., Ninyerola, M., & Pons, X. (2009). Revision of the Single-Channel Algorithm for Land Surface Temperature Retrieval From Landsat Thermal-Infrared Data. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(1), 339–349. *IEEE Transactions on Geoscience and Remote Sensing*. <https://doi.org/10.1109/TGRS.2008.2007125>
- Jouvet, G., Huss, M., Funk, M., & Blatter, H. (2011). Modelling the retreat of Grosser Aletschgletscher, Switzerland, in a changing climate. *Journal of Glaciology*, *57*(206), 1033–1045. <https://doi.org/10.3189/002214311798843359>
- Karimi, N., Farokhnia, A., Karimi, L., Eftekhari, M., & Ghalkhani, H. (2012). Combining optical and thermal remote sensing data for mapping debris-covered glaciers (Alamkouh Glaciers, Iran). *Cold Regions Science and Technology*, *71*, 73–83. <https://doi.org/10.1016/j.coldregions.2011.10.004>
- Kaushik, S., Singh, T., Bhardwaj, A., Joshi, P. K., & Dietz, A. J. (2022). Automated Delineation of Supraglacial Debris Cover Using Deep Learning and Multisource Remote Sensing Data. *Remote Sensing*, *14*(6), 1352. <https://doi.org/10.3390/rs14061352>
- Koetz, B., Bastiaanssen, W., Berger, M., Defournay, P., Del Bello, U., Drusch, M., Drinkwater, M., Duca, R., Fernandez, V., Ghent, D., Guzinski, R., Hoogeveen, J., Hook, S., Lagouarde, J.-P., Lemoine, G., Manolis, I., Martimort, P., Masek, J., Massart, M., ... Sobrino, J. (2018). High Spatio- Temporal Resolution Land Surface Temperature Mission—A Copernicus Candidate Mission in Support of Agricultural Monitoring. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain*, 8160–8162. <https://doi.org/10.1109/IGARSS.2018.8517433>
- Kraaijenbrink, P. D. A., Bierkens, M. F. P., Lutz, A. F., & Immerzeel, W. W. (2017). Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers. *Nature*, *549*(7671), 257–260. <https://doi.org/10.1038/nature23878>
- Lagouarde, J.-P., Bhattacharya, B. K., Crebassol, P., Gamet, P., Babu, S. S., Boulet, G., Briottet, X., Buddhiraju, K. M., Cherchali, S., Dadou, I., Dedieu, G., Gouhier, M., Hagolle, O., Irvine, M., Jacob, F., Kumar, A., Kumar, K. K., Laignel, B., Mallick, K., ... Ramakrishnan, R. (2018). The Indian-French Trishna Mission: Earth Observation in the Thermal Infrared with High Spatio-Temporal Resolution. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 4078–4081. <https://doi.org/10.1109/IGARSS.2018.8518720>
- Langhammer, L., Rabenstein, L., Schmid, L., Bauder, A., Grab, M., Schaer, P., & Maurer, H. (2019). Glacier bed surveying with helicopter-borne dual-polarization ground-penetrating radar. *Journal of Glaciology*, *65*(249), 123–135. <https://doi.org/10.1017/jog.2018.99>
- Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., & Sobrino, J. A. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment*, *131*, 14–37. <https://doi.org/10.1016/j.rse.2012.12.008>
- Li, Z.-L., Wu, H., Duan, S.-B., Zhao, W., Ren, H., Liu, X., Leng, P., Tang, R., Ye, X., Zhu, J., Sun, Y., Si, M., Liu, M., Li, J., Zhang, X., Shang, G., Tang, B.-H., Yan, G., & Zhou, C. (2023). Satellite Remote Sensing of Global Land Surface Temperature: Definition, Methods, Products, and Applications. *Reviews of Geophysics*, *61*(1), e2022RG000777. <https://doi.org/10.1029/2022RG000777>

- Linsbauer, A., Huss, M., Hodel, E., Bauder, A., Fischer, M., Weidmann, Y., Bärtschi, H., & Schmassmann, E. (2021). The New Swiss Glacier Inventory SGI2016: From a Topographical to a Glaciological Dataset. *Frontiers in Earth Science*, 9, 704189. <https://doi.org/10.3389/feart.2021.704189>
- Malakar, N. K., Hulley, G. C., Hook, S. J., Laraby, K., Cook, M., & Schott, J. R. (2018). An Operational Land Surface Temperature Product for Landsat Thermal Data: Methodology and Validation. *IEEE Transactions on Geoscience and Remote Sensing*, 56(10), 5717–5735. [IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2018.2824828](https://doi.org/10.1109/TGRS.2018.2824828)
- Malbêteau, Y., Merlin, O., Gascoin, S., Gastellu, J. P., Mattar, C., Olivera-Guerra, L., Khabba, S., & Jarlan, L. (2017). Normalizing land surface temperature data for elevation and illumination effects in mountainous areas: A case study using ASTER data over a steep-sided valley in Morocco. *Remote Sensing of Environment*, 189, 25–39. <https://doi.org/10.1016/j.rse.2016.11.010>
- Markham, B. L., Storey, J. C., Williams, D. L., & Irons, J. R. (2004). Landsat sensor performance: History and current status. *IEEE Transactions on Geoscience and Remote Sensing*, 42(12), 2691–2694. <https://doi.org/10.1109/TGRS.2004.840720>
- Maussion, F., Hock, R., Paul, F., Raup, B., Rastner, P., Zemp, M., Andreassen, L., Barr, I., Bolch, T., Kochtitzky, W., McNabb, R., & Tielidze, L. (2023). *The Randolph Glacier Inventory version 7.0 User guide v1.0*. <https://doi.org/doi:10.5281/zenodo.8362857>
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*. <https://doi.org/10.1080/01431169608948714>
- Mellor, A., Boukir, S., Haywood, A., & Jones, S. (2015). Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, 155–168. <https://doi.org/10.1016/j.isprsjprs.2015.03.014>
- Mitkari, K. V., Arora, M. K., Tiwari, S. P., Sofat, S., Gusain, H. S., & Tiwari, R. K. (2022). Large-Scale Debris Cover Glacier Mapping Using Multisource Object-Based Image Analysis Approach. *Remote Sensing*, 14(3202). <https://doi.org/10.3390/rs14133202>
- Mölg, N., Bolch, T., Walter, A., & Vieli, A. (2019). Unravelling the evolution of Zmuttgletscher and its debris cover since the end of the Little Ice Age. *The Cryosphere*, 13(7), 1889–1909. <https://doi.org/10.5194/tc-13-1889-2019>
- Mölg, N., Ferguson, J., Bolch, T., & Vieli, A. (2020). On the influence of debris cover on glacier morphology: How high-relief structures evolve from smooth surfaces. *Geomorphology*, 357, 107092. <https://doi.org/10.1016/j.geomorph.2020.107092>
- Nainwal, H. C., Banerjee, A., Shankar, R., Semwal, P., & Sharma, T. (2016). Shrinkage of Satopanth and Bhagirath Kharak Glaciers, India, from 1936 to 2013. *Annals of Glaciology*, 57(71), 131–139. <https://doi.org/10.3189/2016AoG71A015>
- Nakawo, M., Iwata, S., Watanabe, O., & Yoshida, M. (1986). Processes which Distribute Supraglacial Debris on the Khumbu Glacier, Nepal Himalaya. *Annals of Glaciology*, 8, 129–131. <https://doi.org/10.3189/S0260305500001294>
- NASA JPL. (2020). *NASADEM Merged DEM Global 1 arc second V001* [Dataset]. NASA EOSDIS Land Processes DAAC. https://doi.org/doi:10.5067/MEaSURES/NASADEM/NASADEM_HGT.001

- Nunchhani, V., Hazarika, S., Murtem, R., Bandyopadhyay, A., & Bhadra, A. (2024). Analysis of the temporal variations in glaciers' surface area in Alaknanda River Basin, Uttarakhand. *Journal of Water and Climate Change*, 00(0), 2–17. <https://doi.org/10.2166/wcc.2024.593>
- Paul, F., Kääb, A., Maisch, M., Kellenberger, T., & Haeberli, W. (2004). Rapid disintegration of Alpine glaciers observed with satellite data. *Geophysical Research Letters*, 31(21), 2004GL020816. <https://doi.org/10.1029/2004GL020816>
- Peña-Haro, S., Lukes, R., Carrel, M., & Lüthi, B. (2021). Image-based flow measurements in wide rivers using a multi-view approach. *Interpraevent Conference Proceedings*, 146–153.
- Pope, A., & Rees, G. (2014). Using *in situ* spectra to explore Landsat classification of glacier surfaces. *International Journal of Applied Earth Observation and Geoinformation*, 27, 42–52. <https://doi.org/10.1016/j.jag.2013.08.007>
- Qiu, S., Zhu, Z., Olofsson, P., Woodcock, C. E., & Jin, S. (2023). Evaluation of Landsat image compositing algorithms. *Remote Sensing of Environment*, 285, 113375. <https://doi.org/10.1016/j.rse.2022.113375>
- Ranzi, R., Grossi, G., Iacovelli, L., & Taschner, S. (2004). Use of multispectral ASTER images for mapping debris-covered glaciers within the GLIMS Project. In *International Geoscience and Remote Sensing Symposium (IGARSS)* (Vol. 2, p. 1147). <https://doi.org/10.1109/IGARSS.2004.1368616>
- Raper, S. C. B., & Braithwaite, R. J. (2006). Low sea level rise projections from mountain glaciers and icecaps under global warming. *Nature*, 439(7074), 311–313. <https://doi.org/10.1038/nature04448>
- RGI 7.0 Consortium. (2023). *Randolph Glacier Inventory—A Dataset of Global Glacier Outlines* (Version Version 7.0) [Dataset]. Boulder; National Snow and Ice Data Center. <https://doi.org/10.5067/f6jmovy5navz>
- Salvatore, M. C., Zanoner, T., Baroni, C., Carton, A., Banchieri, F. A., Viani, C., Giardino, M., & Perotti, L. (2015). The state of Italian glaciers: A snapshot of the 2006–2007 hydrological period. *Geografia Fisica e Dinamica Quaternaria*, 38(2), 175–198.
- Scherler, D., Bookhagen, B., & Strecker, M. R. (2011a). Hillslope-glacier coupling: The interplay of topography and glacial dynamics in High Asia. *Journal of Geophysical Research: Earth Surface*, 116(F2). <https://doi.org/10.1029/2010JF001751>
- Scherler, D., Bookhagen, B., & Strecker, M. R. (2011b). Spatially variable response of Himalayan glaciers to climate change affected by debris cover. *Nature Geoscience*, 4(3), 156–159. <https://doi.org/10.1038/ngeo1068>
- Schimel, D. S., & Poulter, B. (2022). The Earth in Living Color—NASA's Surface Biology and Geology Designated Observable. *2022 IEEE Aerospace Conference (AERO)*, 1–6. <https://doi.org/10.1109/AERO53065.2022.9843640>
- Schmidt, G., Jenkerson, C. B., Masek, J., Vermote, E., & Gao, F. (2013). Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description. In *Open-File Report* (Nos. 2013–1057). U.S. Geological Survey. <https://doi.org/10.3133/ofr20131057>
- Shah, S. S., Banerjee, A., Nainwal, H. C., & Shankar, R. (2019). Estimation of the total sub-debris ablation from point-scale ablation data on a debris-covered glacier. *Journal of Glaciology*, 65(253), 759–769. <https://doi.org/10.1017/jog.2019.48>

Sharda, S., & Srivastava, M. (2024). Mapping of Debris-Covered Glaciers Using Object-Based Machine Learning Technique. *Journal of the Indian Society of Remote Sensing*. <https://doi.org/10.1007/s12524-024-01832-2>

Shukla, A., Gupta, R. P., & Arora, M. K. (2010). Delineation of debris-covered glacier boundaries using optical and thermal remote sensing data. *Remote Sensing Letters*, 1(1), 11–17. <https://doi.org/10.1080/01431160903159316>

Skakun, S., Vermote, E. F., Roger, J.-C., Justice, C. O., & Masek, J. G. (2019). Validation of the LaSRC Cloud Detection Algorithm for Landsat 8 Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7), 2439–2446. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. <https://doi.org/10.1109/JSTARS.2019.2894553>

Taloor, A. K., Manhas, D. S., & Kothyari, G. C. (2021). Retrieval of land surface temperature, normalized difference moisture index, normalized difference water index of the Ravi basin using Landsat data. *Applied Computing and Geosciences*, 9, 100051. <https://doi.org/10.1016/j.acags.2020.100051>

Taschner, S., & Ranzi, R. (2002). Comparing the opportunities of Landsat-TM and Aster data for monitoring a debris covered glacier in the Italian Alps within the GLIMS project. *IEEE International Geoscience and Remote Sensing Symposium*, 2, 1044–1046. <https://doi.org/10.1109/IGARSS.2002.1025770>

USGS. (2021). Landsat Collection 2 Level-2 Science Products. In *Fact Sheet* (Nos. 2021–3055). U.S. Geological Survey. <https://doi.org/10.3133/fs20213055>

Vincent, O., & Folorunso, O. (2009). *A Descriptive Algorithm for Sobel Image Edge Detection*. InSITE 2009: Informing Science + IT Education Conference. <https://doi.org/10.28945/3351>

Xu, X., Xu, S., Jin, L., & Song, E. (2011). Characteristic analysis of Otsu threshold and its applications. *Pattern Recognition Letters*, 32(7), 956–961. <https://doi.org/10.1016/j.patrec.2011.01.021>

Zeller, L., McGrath, D., & McCoy, S. W. (2022). *Supraglacial lakes on debris-covered glaciers: High resolution observations of subseasonal lake expansion, drainage, and the controlling processes*. C52E-0400. AGU Fall Meeting Abstracts.

Appendix

A Google Earth Engine Scripts

The scripts can be accessed via Google Earth Engine (GEE), or via a Github repository.

Access to the complete GEE repository:

https://code.earthengine.google.com/?accept_repo=users/lorenamueller/Refine_DCG_Outlines

The individual scripts can be accessed in GEE using the following links:

1. Data selection and generation of input layers :
<https://code.earthengine.google.com/c9ec0115fe6a3a9cf7caba6d3d026777>
2. DCG classification
 - 2.1 Temporal change analysis :
<https://code.earthengine.google.com/edc08407b2806461cf9ed1f6c3236dac>
 - 2.2 Cross-glacier analysis Approach I :
<https://code.earthengine.google.com/ec3a2c3f327361dec06de5abec84afdd>
 - 2.3 Cross-glacier analysis Approach II :
<https://code.earthengine.google.com/f99ef368609dc8816f57b6391595de75>
 - 2.4 Cross-glacier analysis Approach III :
<https://code.earthengine.google.com/42f0f4394a22b95070c2a2f69a2afe96>

For a complete workflow, first the data selection and generation of input layers must be run, either based on manually selected single scenes or on the creation of a composite image (can be selected inside the script). The generated layers are then used for classification with any of the introduced approaches in a next step.

Github repository: https://github.com/LorenaMueller/refine_DCG_outlines

B Classification Results

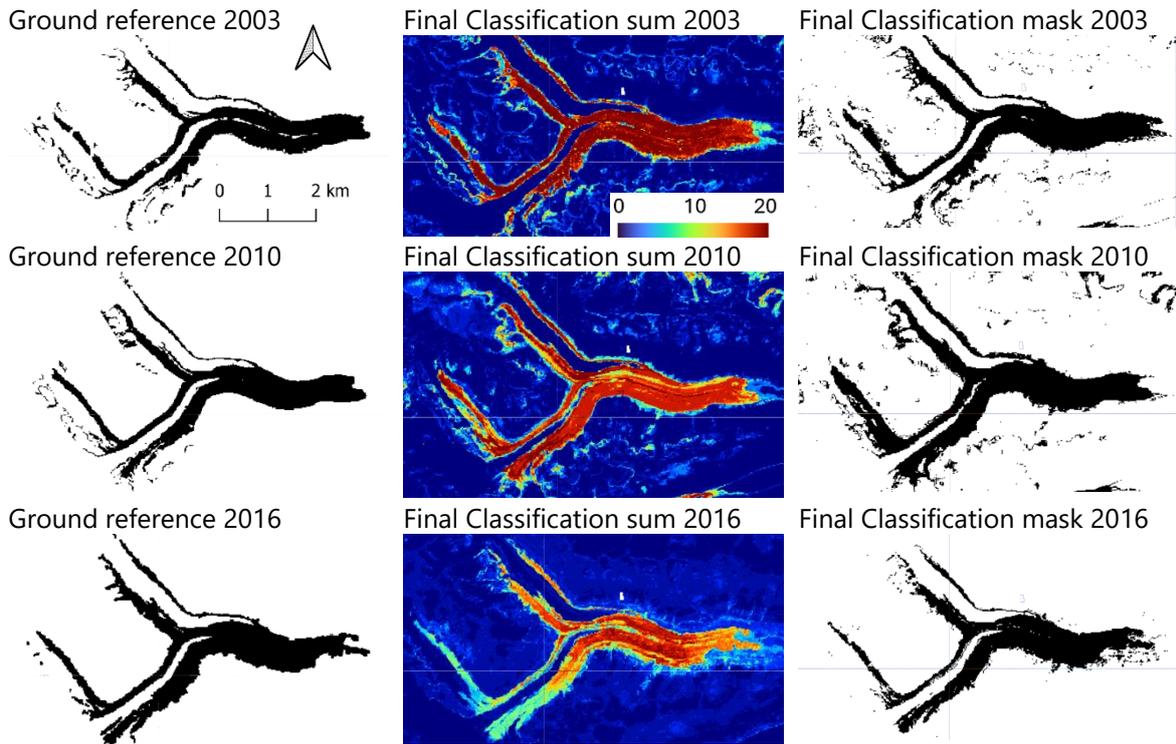


Figure B 1: Classification results of Temporal changes, Unteraargletscher. The colorbar in the final classification sum shows the number of times a pixel was classified as DCG by the different band combinations.

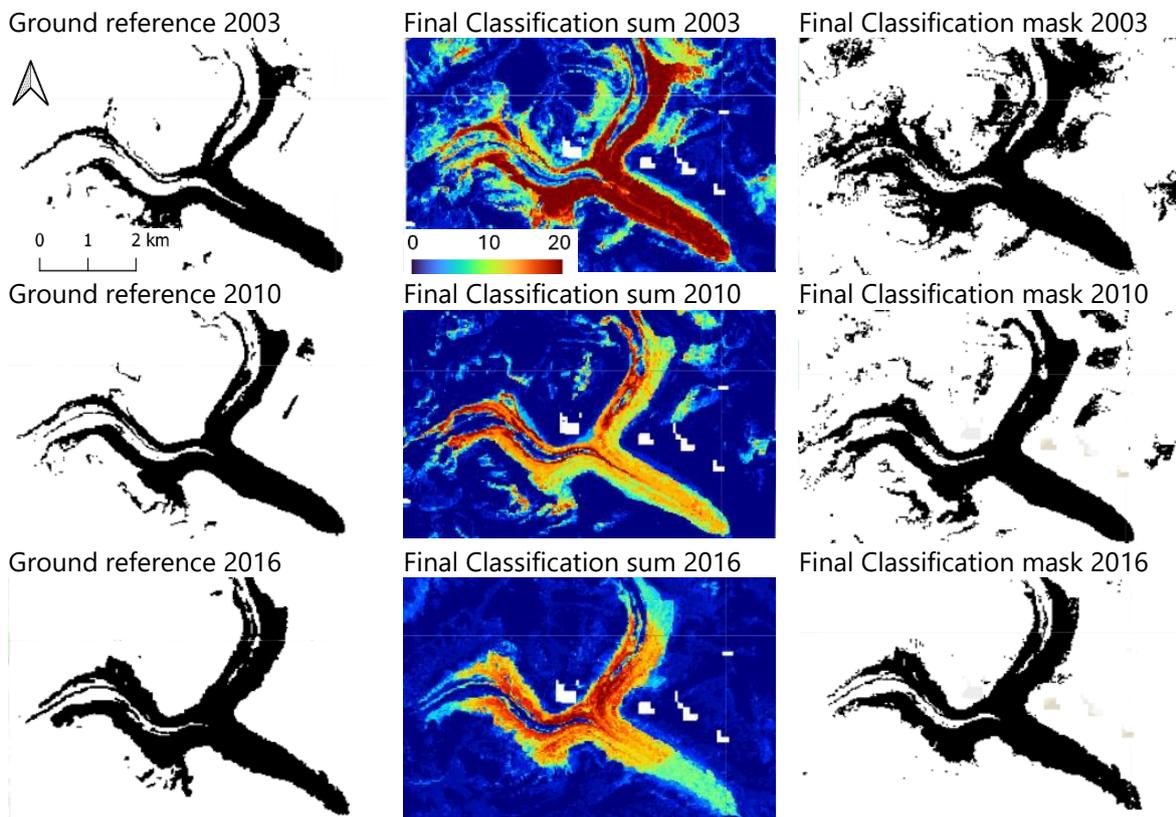


Figure B 2: Classification results of Temporal changes, Oberaletschgletscher.

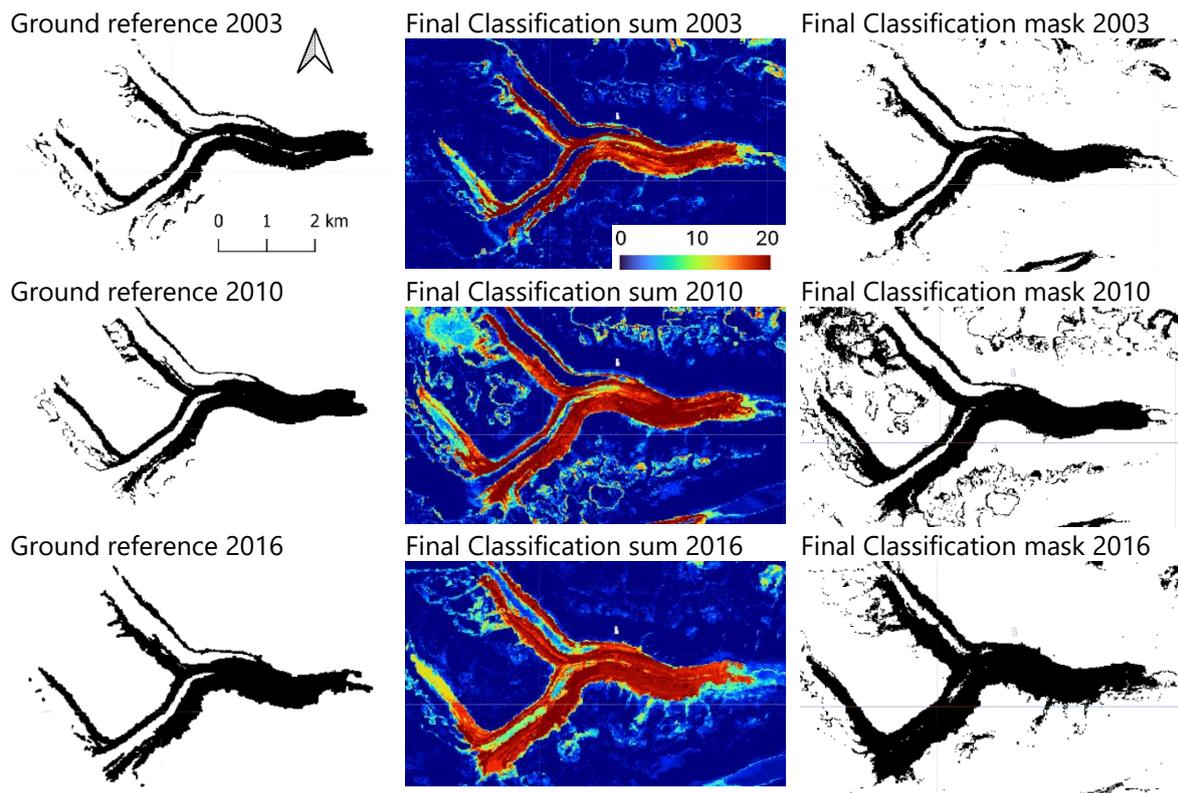


Figure B 3: Classification results of Cross-glacier analysis, Approach I, Unteraargletscher.

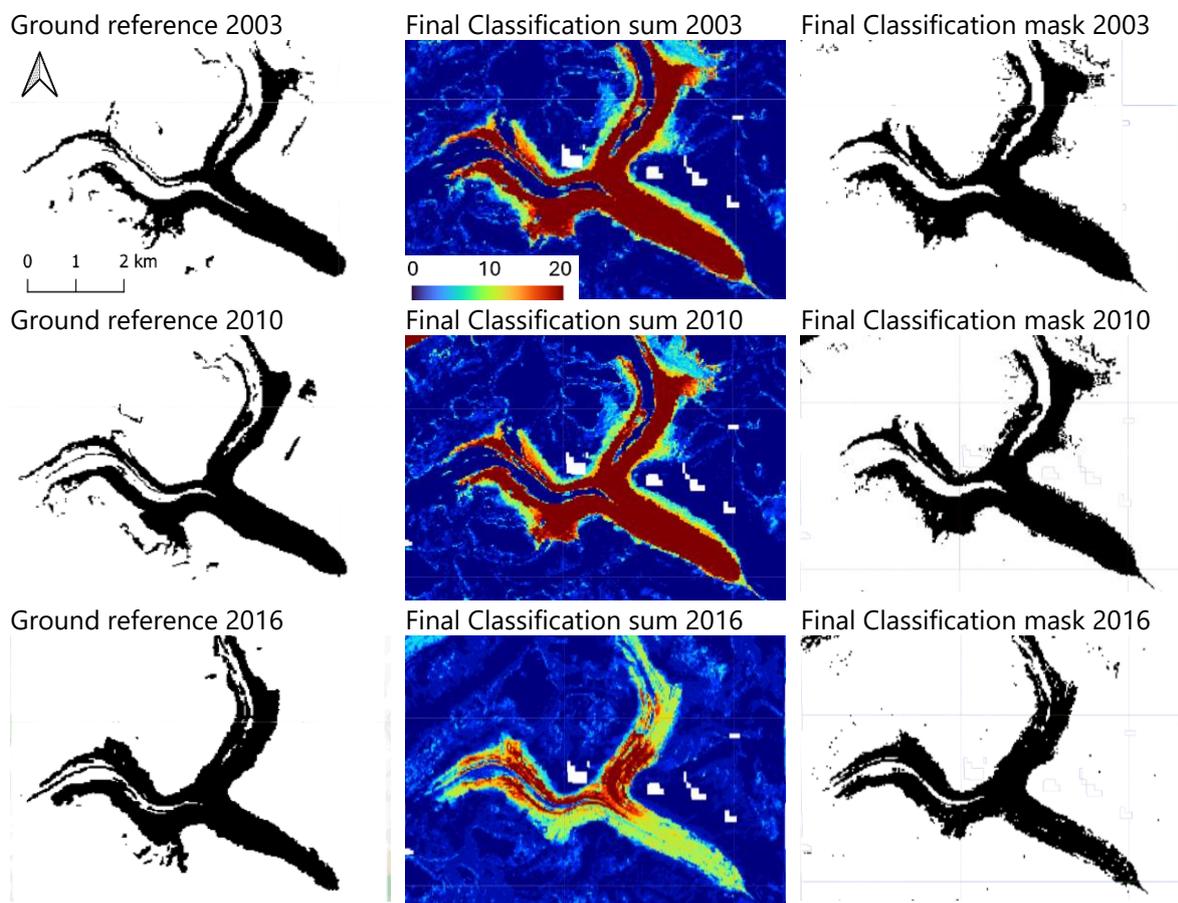


Figure B 4: Classification results of Cross-glacier analysis, Approach I, Oberaletschgletscher.

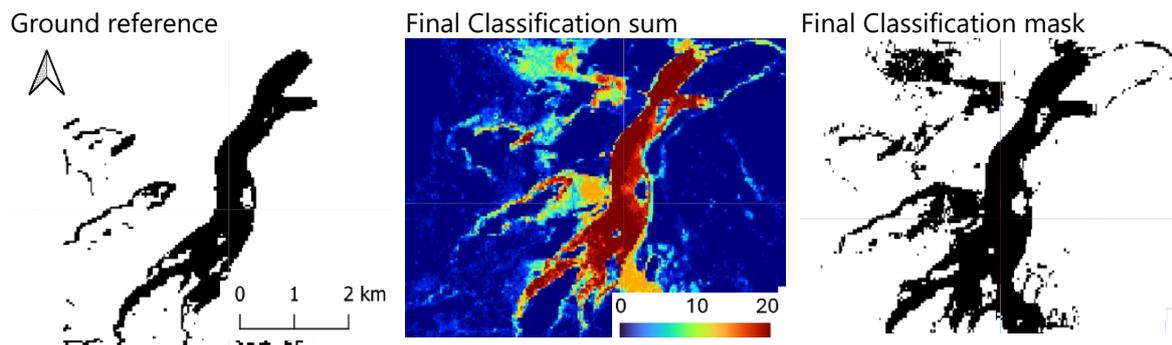


Figure B 5: Classification results of Cross-glacier analysis, Approach I, Belvedere glacier.

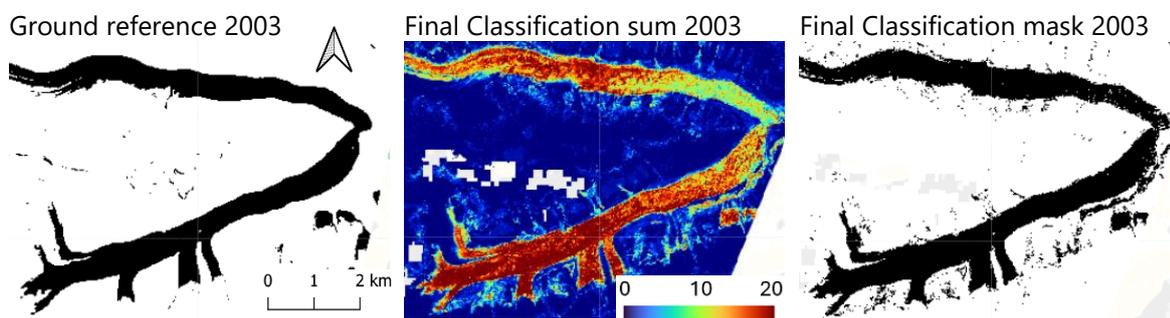


Figure B 6: Classification results of Cross-glacier analysis, Approach I, Satopanth glacier

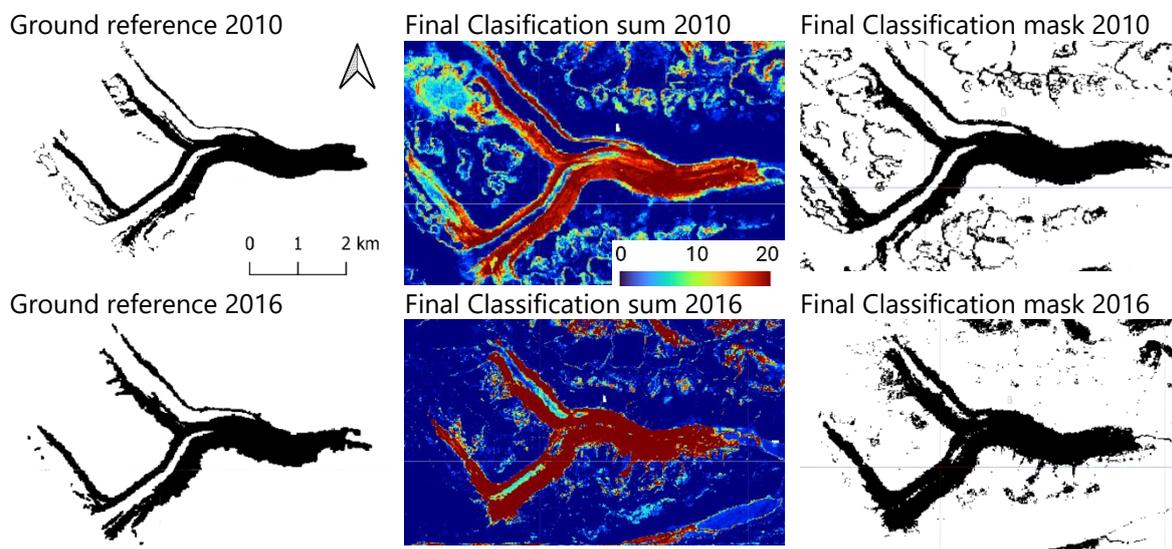


Figure B 7: Classification results of Cross-glacier analysis, Approach II, Unteraargletscher.

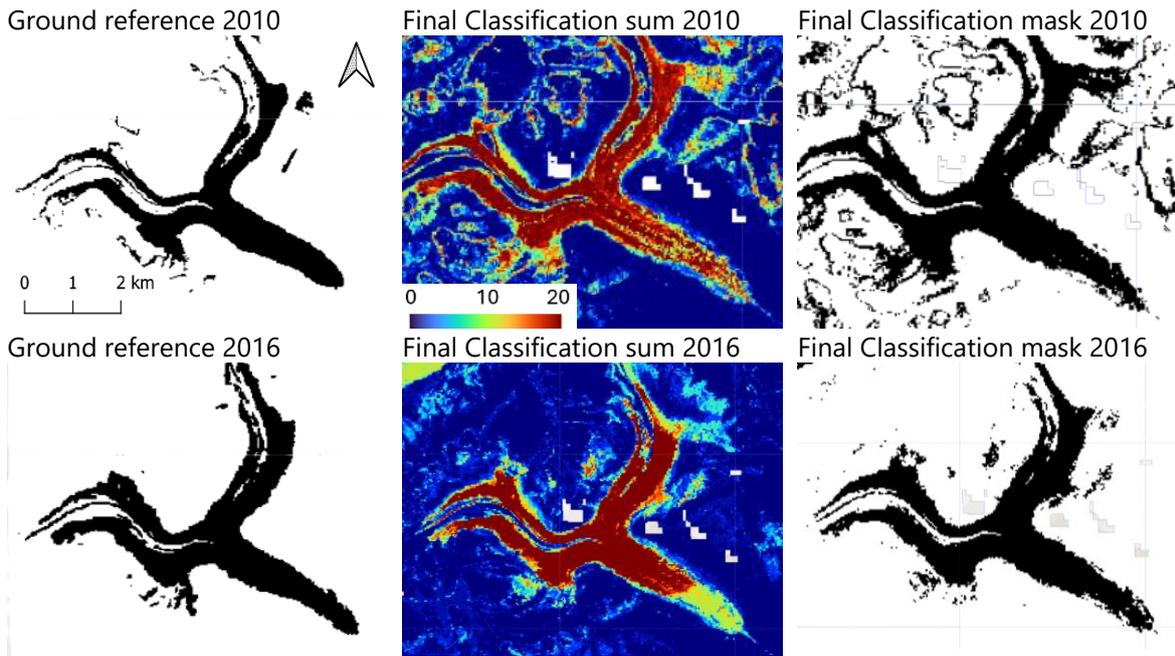


Figure B 8: Classification results of Cross-glacier analysis, Approach II, Oberaletschgletscher.

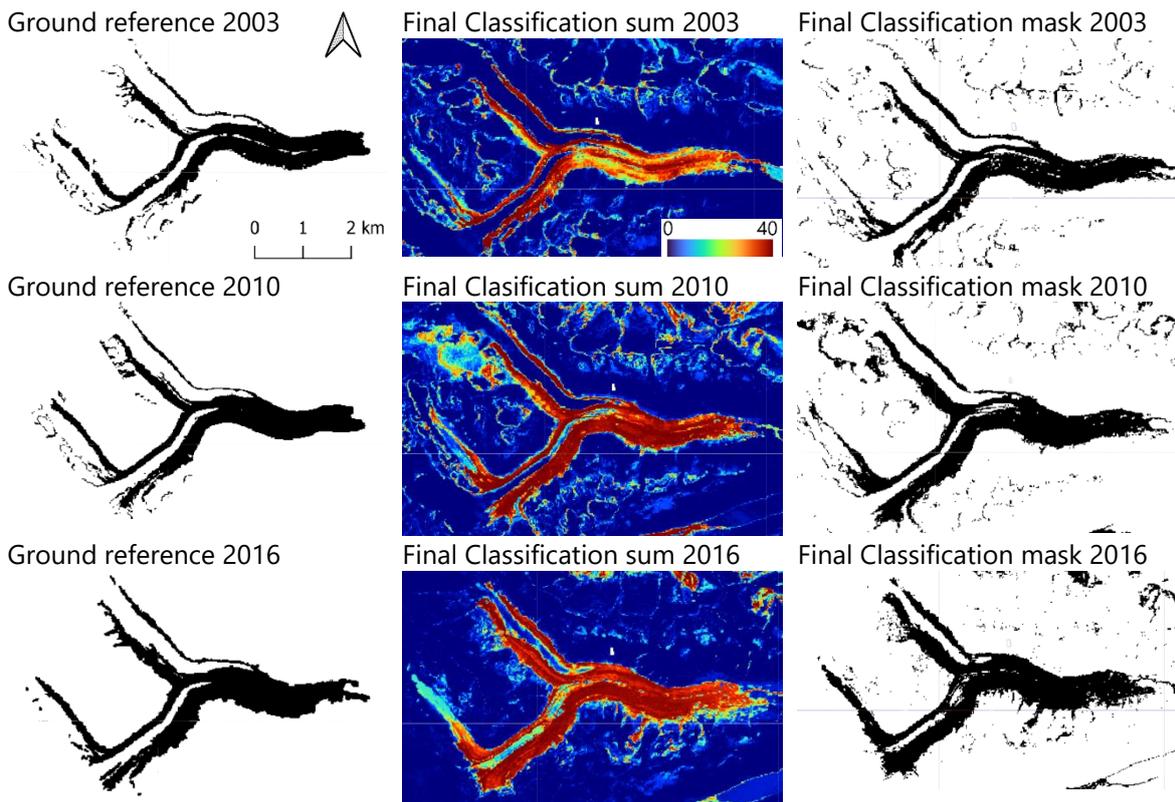


Figure B 9: Classification results of Cross-glacier analysis, Approach III, Unteraargletscher.

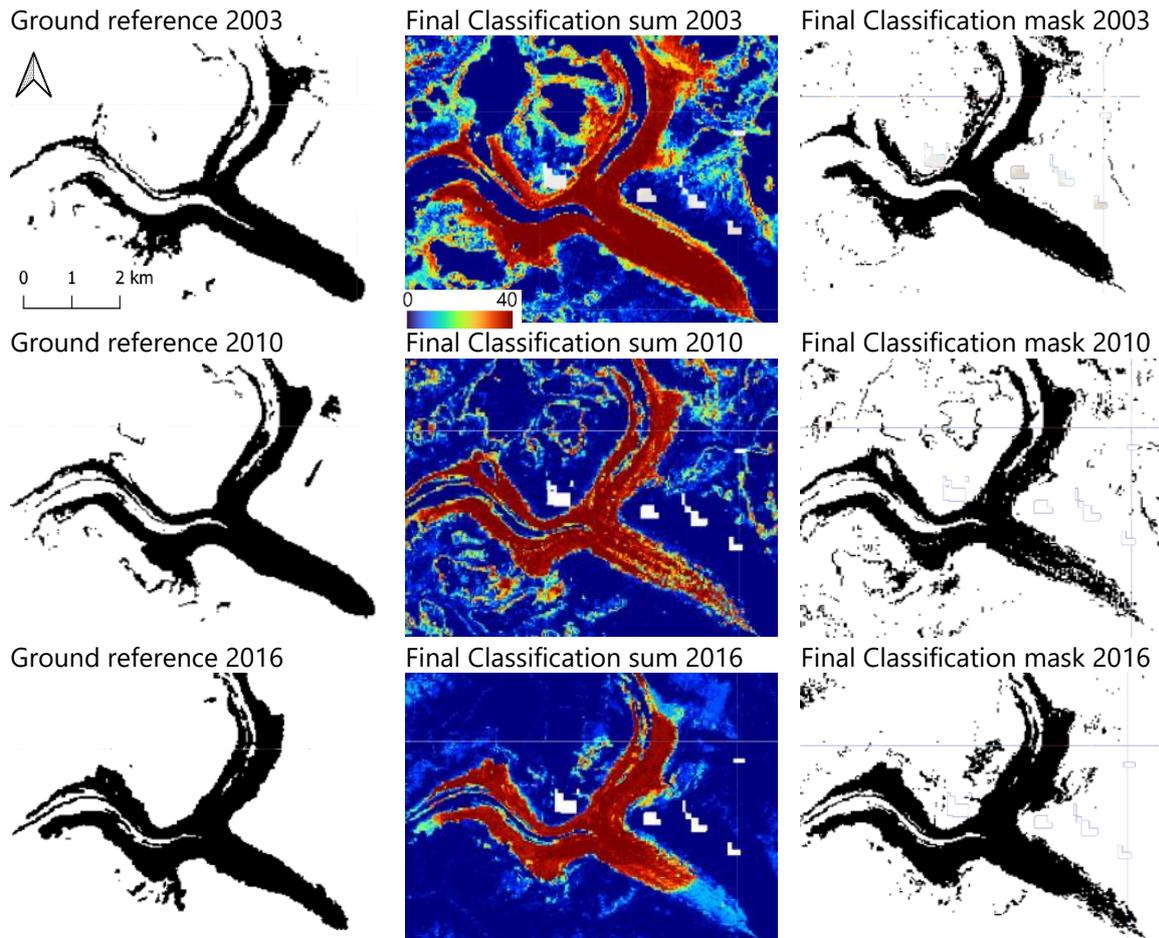


Figure B 10: Classification results of Cross-glacier analysis, Approach III, Oberaletschgletscher.

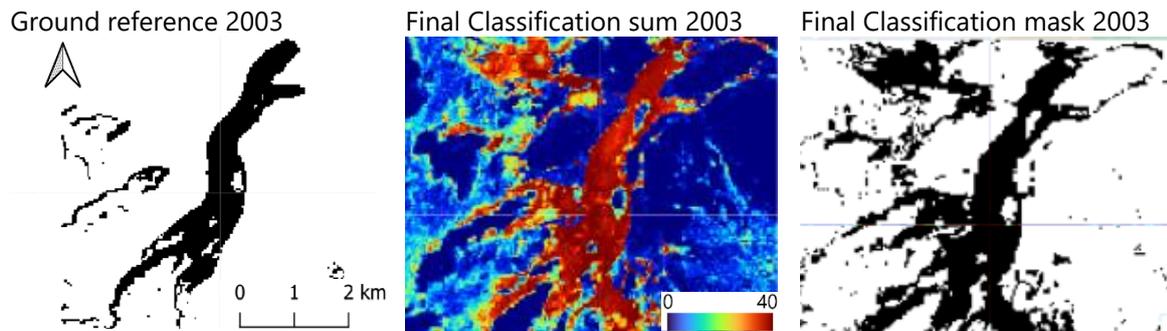


Figure B 11: Classification results of Cross-glacier analysis, Approach III, Belvedere glacier.

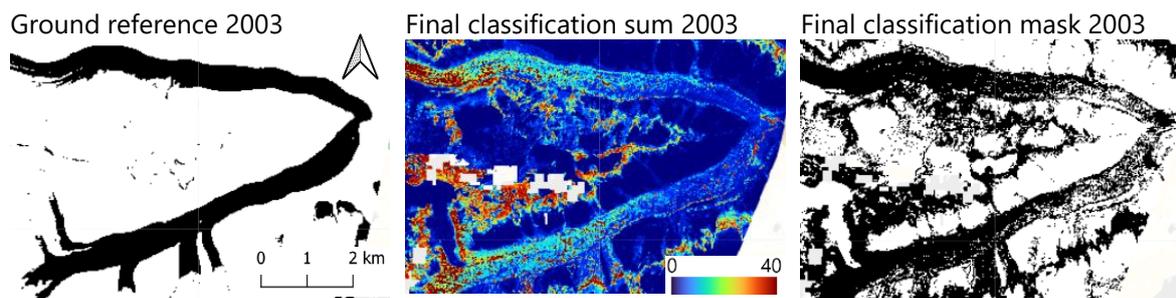
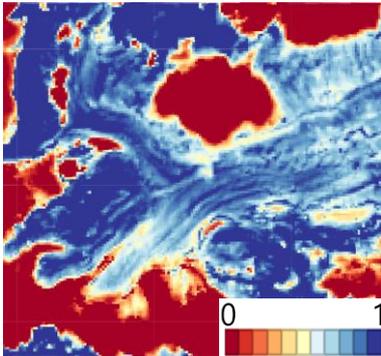


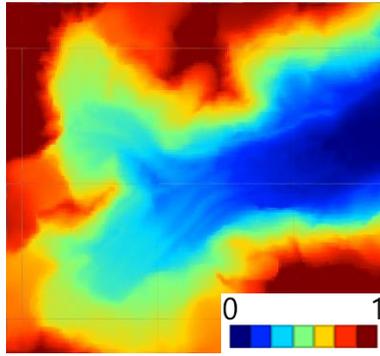
Figure B 12: Classification results of Cross-glacier analysis, Approach III, Satopanth glacier.

C Input Layers

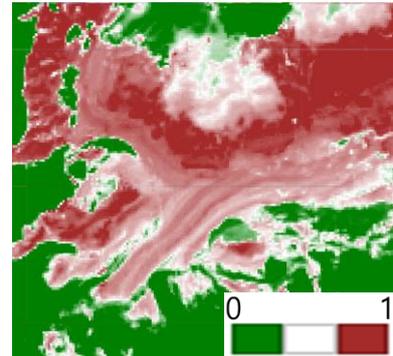
LST NIR Index



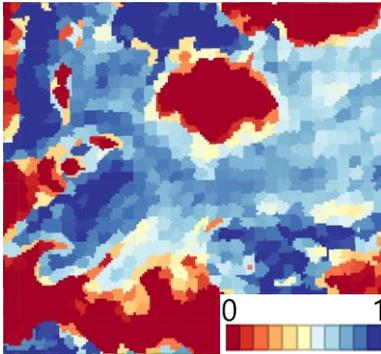
Elevation



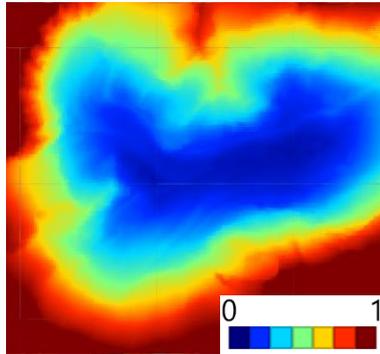
NDRI



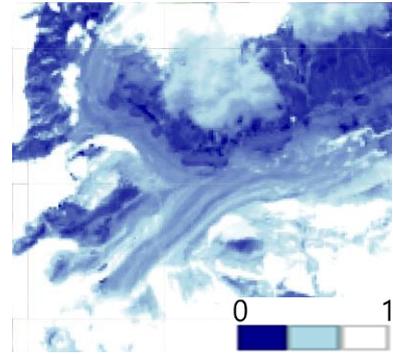
LST NIR superpixel means



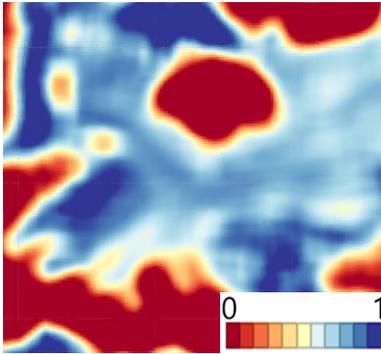
TPI



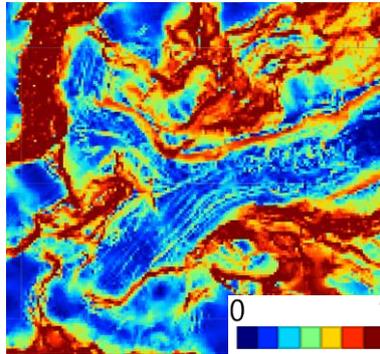
NDSI



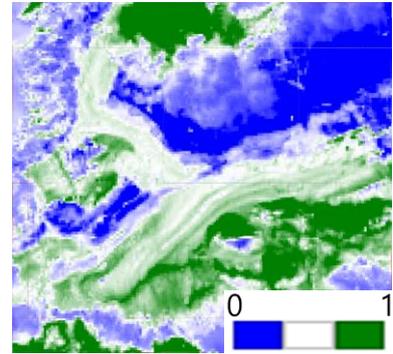
LST NIR smooth



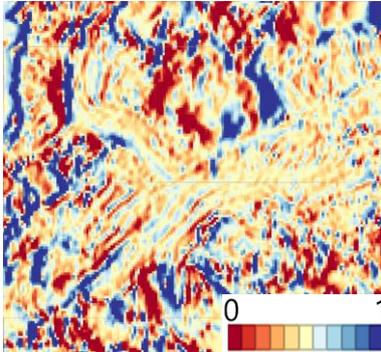
Slope



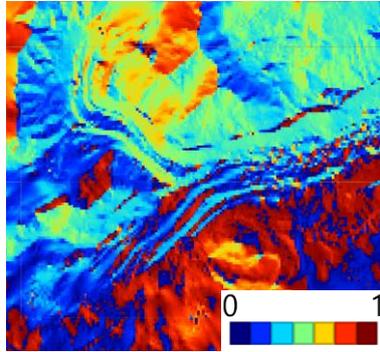
NDWI



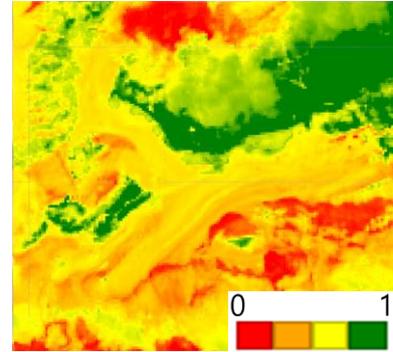
LST NIR edges



Aspect



NDVI



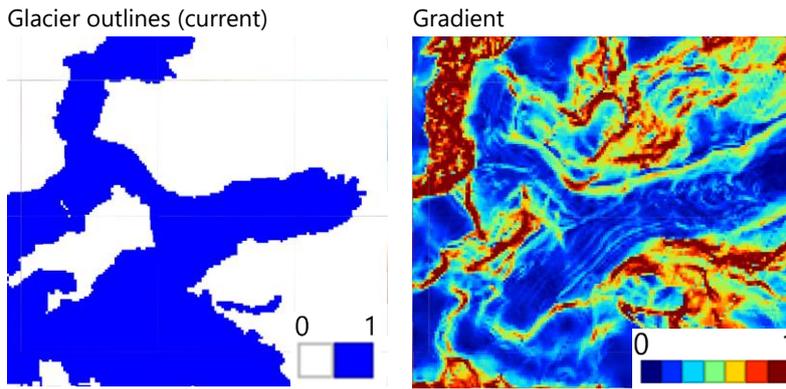


Figure C 1: Overview of the input layers for the Random Forest Classification, for Zmuttgletscher 2016. All input layers were scaled to a value range of [0, 1]. The inclusion of Glacier outlines (current) as displayed here corresponds to the input dataset of temporal change analysis and cross-glacier analysis Approach I.

D Ground reference for 2003 and 2010

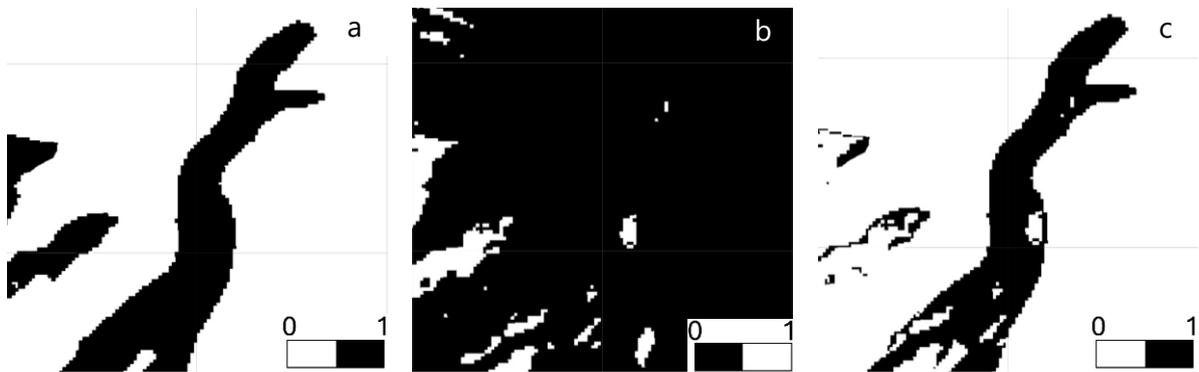


Figure D 1: Creation of ground reference exemplary for Belvedere glacier 2003, with (a) mask of RGI 7.0 glacier outlines (with 0 = no glacier and 1 = glacier), (b) mask of the Normalised Difference Snow Index (NDSI) after applying a threshold of 0.4 (with 0 = no snow/ice and 1 = snow/ice), and (c) resulting ground reference after subtracting the NDSI mask from the RGI outlines (with 0 = no DCG and 1 = DCG).

E Cross-comparison of normalised difference indices with LST

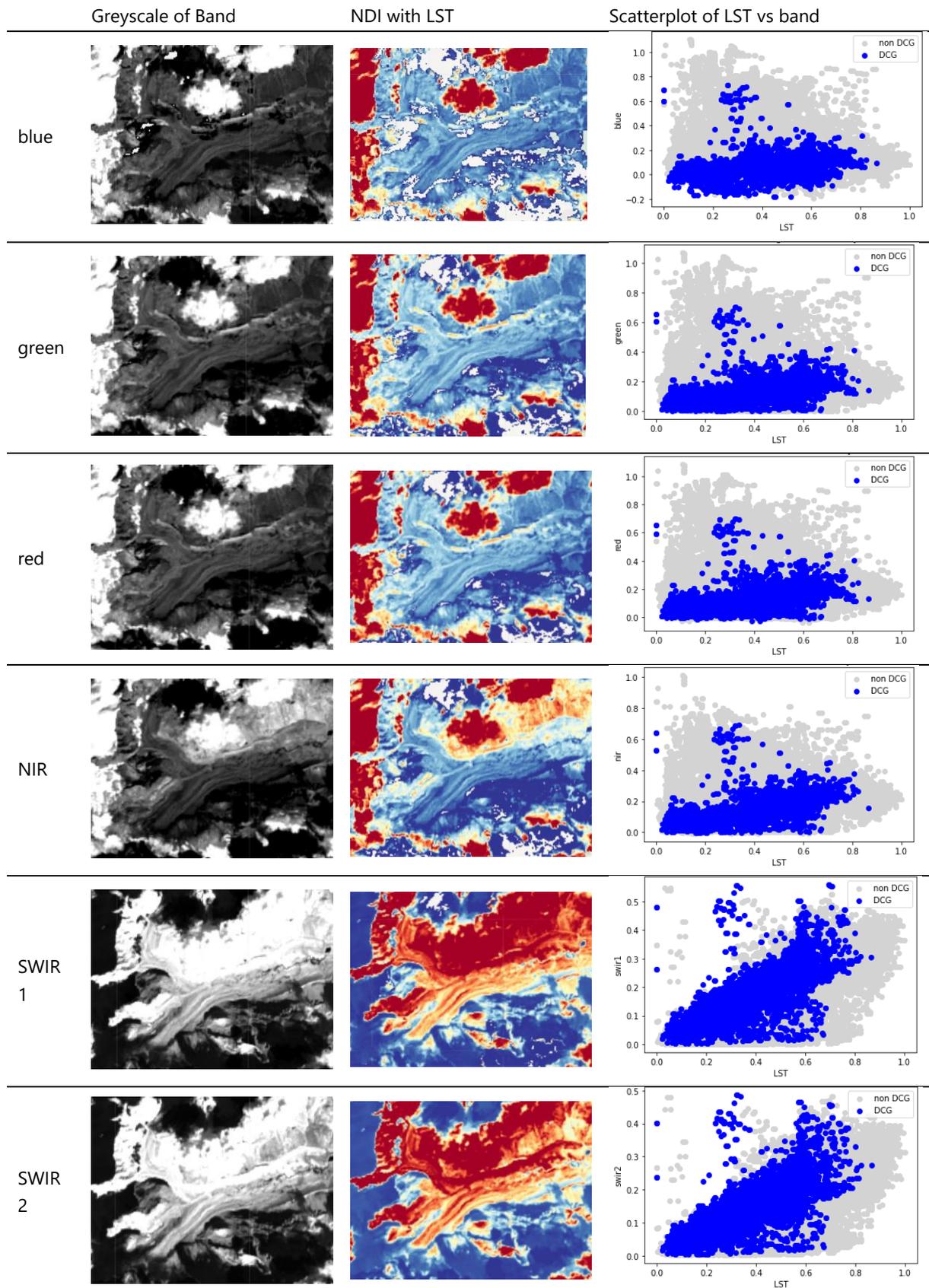


Figure E 1: Complete comparison of the bands tested in a normalised difference index (NDI) with LST, showing the greyscales of each band, the normalised difference index with LST, and a scatterplot of the relation between the respective band (y-axis) and LST (x-axis), with data points coloured as non-DCG (grey) and DCG (blue), according to the SGI2016 debris cover outlines. 91

Personal Declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Date

December 20, 2024

Signature

Lorena Müller

A handwritten signature in black ink, appearing to read 'L. Müller', written in a cursive style.