https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Automated glacial lake extraction using an Object-Based Image Analysis approach in Google Earth Engine

Tomos Morgan, Robert W. McNabb, Paul Dunlop
School of Geography and Environmental Sciences, Ulster University, Coleraine, UK

Correspondence to: Tomos Morgan (Morgan-T15@ulster.ac.uk)

Abstract

The combination of glacial retreat and climate change is increasing the number and size of glacial lakes globally. Many of these glacial lakes are in dangerous glaciated environments, and satellite remote sensing provides a way to improve monitoring efforts, though automated methods are needed to accurately and rapidly detect changes in these lakes. We undertake a total of 40 classification experiments to investigate the impact of classifier parameters, input features and training data on classification accuracy. We run 18 additional experiments to identify the optimal combination of Simple Non-Iterative Clustering segmentation parameters (connectivity and *neighborhoodSize*), assess the impact of input features, determine the required number of training and testing images and compare water extraction indices for the OBIA classification. Our results show that the best-performing combination of parameters was 100-250 training points per class, and values of four and 128 for connectivity and *neighborhoodSize*, respectively. The inclusion of input features such as hillshade, slope, the NDVI and MNDWI in our OBIA classifier improves the overall delineation of glacial lakes and other land classes in our study, particularly in shadow bodies, which are commonly misclassified as water bodies. Finally, we demonstrate that it is possible to accurately classify a time series of images using a single training image, with superior results compared to training with multiple images. We hypothesise that this is due to the complexities of radiometric sensitivity, heterogenous values for bands and indices and temporal changes in land cover throughout the study. Our OBIA approach is a more efficient and accurate way in mapping glacial lakes using Landsat 4-9 satellite imagery over traditional pixel-based approaches, with an overall accuracy of 94.6%, with a producer's accuracy and user's accuracy of 95.3% and 95.5% respectively, for water. This suggests that this method has the potential to map glacial lakes accurately and rapidly over larger regions.

1 Introduction

Glacial lakes are increasing globally, with estimates that global glacier lake volume increased by around 48%, to 156.5 km^3, between 1990-2018, driven by climatic changes that are increasing ablation rates, ice velocities and glacial retreat (Shugar et al., 2020). Many of these glacial lakes are in less accessible and dangerous glaciated environments, and therefore satellite remote sensing has been used to study glacial lakes over a sustained period (Nie et al., 2018; Zhang et al., 2022). Several techniques have been used for lake and glacial lake extraction using satellite remote sensing, including manual delineation (e.g., Wang et al., 2015); Pixel-Based Image Analysis (PBIA; Lillesand et al., 2015), Object-Based Image Analysis (OBIA; Blaschke, 2010), and thresholding based on indices such as the Normalized Difference Water Index (NDWI; McFeeters, 1996), the Modified Normalized Difference Water Index (MNDWI; Xu, 2006) the combination of the MNDWI and Normalized Difference Snow (NDSI; Hall et al., 1995), or the Automated Water Extraction Index (AWEI; Feyisa et al., 2014). Manual

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

delineation is often seen as one of the most accurate ways of monitoring and documenting the areal changes in glacial lakes, but it is time-consuming and may produce errors based on the image resolution, operator bias and experience (Hall et al., 2003; Tang et al., 2024; Wang et al., 2015).

Pixel-based classification approaches utilise the spectral information of each pixel to help determine the type(s) of landcover within that pixel. Frequently used examples for mapping glacial lakes include the NDWI, MNDWI or other satellite-derived indices to map water bodies and other landcover types (e.g., Li and Sheng, 2012; Shugar et al., 2020; Wang et al., 2020; Yan et al., 2020; Zhang et al., 2022). The MNDWI has been found to perform better than the NDWI in many regions, with relatively stable thresholds in comparison with other water indices (Ji et al., 2009; Feyisa et al., 2014; Zhang et al., 2015; Singh et al., 2015; Zhao et al., 2018). Recent studies have utilised the MNDWI index using a threshold of 0 that produced an overall accuracy of higher than 90% in glacial lake mapping in High Mountain Asia in 2015 from Landsat 8 imagery (Tian et al., 2017; Zhao et al., 2018). The AWEI_sh can extract water in mountainous areas with high accuracy, avoiding classifying deep shadow as water which is a common classification error in many indices (Feyisa et al., 2014).

Despite successes in using pixel-based indices to detect changes in glacial lakes, results are highly variable and each region typically requires a different threshold value to detect water due to physical properties such as mountain, cloud cover, steep terrain, seasonally frozen lake water, lake turbidity, iceberg calving and melting glaciers (Bhambri et al., 2015; Jawak and Luis, 2014; Li and Sheng, 2012; Mitkari et al., 2017; Qayyum et al., 2020; Rastner et al., 2014; Sahu and Gupta, 2018; Wang et al., 2013; Wessels et al., 2002; Xie et al., 2013). This is a further limitation to using a PBIA approach for large-scale mapping, as manual inspection of threshold values for multiple lakes, images and satellite sensors is time-consuming. Additionally, applying the MNDWI and the AWEI_sh or AWEI_nsh without a decision tree to an image with high albedo surfaces may lead to large false positive values and is documented to be unable to reject snow cover and shadow, due to the absence of the support of the machine learning classification algorithm that aids in categorising pixels based on their spectral values (Acharya et al., 2018; Feyisa et al., 2014; Yang et al., 2003). A decision tree is defined as a basic concept of splitting a complex decision into several simpler decisions, where features of data (bands) are predictor variables and the classes to be mapped are referred to as the target variables (Breiman et al., 2017; Xu et al., 2005). Using a classification decision tree is advantageous because it can handle complex nonlinear relationships between spectral bands and class proportions of mixed pixels in images, utilising the terrain, texture, spectral and spatial distributions to enhance the classification accuracy (Xu et al., 2005; Yang et al., 2017).

There is a need for an automated method that can accurately monitor and map glacial lakes at the regional scale, to have a clearer understanding of their future trajectory without the need for extensive manual inspection or correction. The alternative to a pixel-based approach is applying an OBIA approach that deals with image segments (a group of pixels) instead of individual pixels, enabling the inclusion of both spatial and spectral features, which is known to improve the accuracy of the resultant classified map (Blaschke, 2010; Capolupo et al., 2020; Castilla and Hay, 2008; Hossain and Chen, 2019; Whiteside et al., 2011). One of the most applied classification algorithms is Random Forest (RF), defined as an ensemble classifier that produces multiple decision trees that “vote” on segments or groups and then classifies them based on the most popular class using randomly selected subsets of training samples and variables (Belgiu and Drăguţ, 2016; Lu et al., 2021; Pal, 2005).

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

The two basic principles of OBIA are segmentation and classification. The segmentation of an image is defined as the process of partitioning an image into different segments or groups, which are then classified based on their geometry, area, colour, shape and texture (Su and Zhang, 2017). In the context of mapping glacial lakes, the main advantages of using OBIA over pixel-based approaches are that it can map glacial lakes with higher accuracy and minimise the task of automatic corrections (Mitkari et al., 2017; Su and Zhang, 2017). A recent study conveyed its superiority on the overall detection of different classes (snow/ice, ice-mixed debris, supraglacial debris, periglacial debris, valley rock, glacial lakes, exposed ice-faces, crevasses, rills, debris cones and meltwater streams) on the Gangotri Glacier, India using both a PBIA and OBIA approach. The study showed that the OBIA was superior with an overall accuracy of 93.6% compared to 89.1% for a PBIA approach (Mitkari et al., 2022).

Mapping water bodies over a large study area and over decades using multiple satellite images is data intensive, with file sizes ranging from ~200 MB for complete Landsat 4 and 5 scenes to ~1 GB for Landsat 8 and 9 scenes (Ali et al., 2023). Developing an approach using the cloud-based capabilities of Google Earth Engine (GEE; Gorelick et al., 2017) enables users to reduce the time and cost of downloading, storing and processing images locally, allowing this approach to be quickly applied to the whole Landsat collection for the chosen study area. In this study we present an OBIA approach to mapping glacial lakes developed using GEE. We apply this approach to various Landsat scenes, testing the impact of varying segmentation parameters and input features.

Our aims are:

1) To determine the most efficient automated method to classify water by conducting experiments within GEE to investigate the effects of changing different segmentation and classifier parameters.
2) To investigate the importance that the inclusion of input features such as hillshade, slope, MNDWI and NDVI has in the classification of water and other landcover types.
3) To study whether training a classifier with a single image can produce accurate classification maps when applied across the whole Landsat collection.
4) To compare the training of a single image against one image per Landsat sensor (five in total) to assess the changes in overall accuracies of detecting various landcover types in the OBIA classifier.
5) To produce a detailed accuracy assessment comparison between different water delineation methods and the OBIA to assess which method(s) is most efficient in extracting glacial lakes in an image collection.
6) To assess whether the OBIA classifier produced in this study can outperform pixel-based indices.

1.1 Study area

New Zealand consists of two main islands, the North and South, with most glaciers residing on the South Island in the Southern Alps. The mountain range is 700 km long and ~50 km wide, beginning at 43.5°S to 170.5°E and peaking at 3724 a.s.l at Aoraki / Mount Cook that formed due to compression of the Pacific and Australian plates (Garcia, 2022; Spasojević and Clayton, 2008; Fig.1; Sturman and Tapper, 1996). The mountain range acts as a barrier to the prevailing westerly atmospheric circulation from the Tasman Sea, with the central divide between the east and west of the Southern Alps receiving up to 15 m a^{-1} of mean annual precipitation, which drops to ~1.5 m a^{-1} at the southeast of the main drainage divide (Carrivick et al., 2022; Henderson and Thompson, 1999; Paul et al., 2023). Frequent cloud cover is derived from steep precipitation gradients in this region that provides protection from direct solar radiation (Baumann et al., 2021; Chinn, 2001; Paul et al., 2023).

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Most of the research on lake-terminating glaciers in New Zealand has been heavily weighted towards the Tasman glacier that is 23km long, terminating in a large proglacial lake and lying east of the main divide of the Southern Alps (Dykes, 2013). The Tasman glacier and lake is well-documented in literature with various studies investigating and identifying phases of glacial retreat, calving rates and changes in glacial lake area (Baumann et al., 2021; Carrivick et al., 2022; Chinn et al., 2012, 2005; Dykes, 2013; Dykes et al., 2011; Garcia, 2022; Hochstein et al., 1995; Mackintosh et al., 2017; Robertson et al., 2012; Röhl, 2006, 2008; Warren and Kirkbride, 1998). A recent study produced by Carrivick et al. (2022) used a combination of historical measurements and manually digitised lake outlines for all ice-marginal lakes in New Zealand. This study reported an increase in ice-marginal lake area of 159.8%, from 9.2 km^2 in 1990 to 23.9 km^2 in 2020 with model estimations of a further increase in combined size by ~ 150% towards 2050, followed by a decrease to 2100 as glaciers disconnect from their lakes.

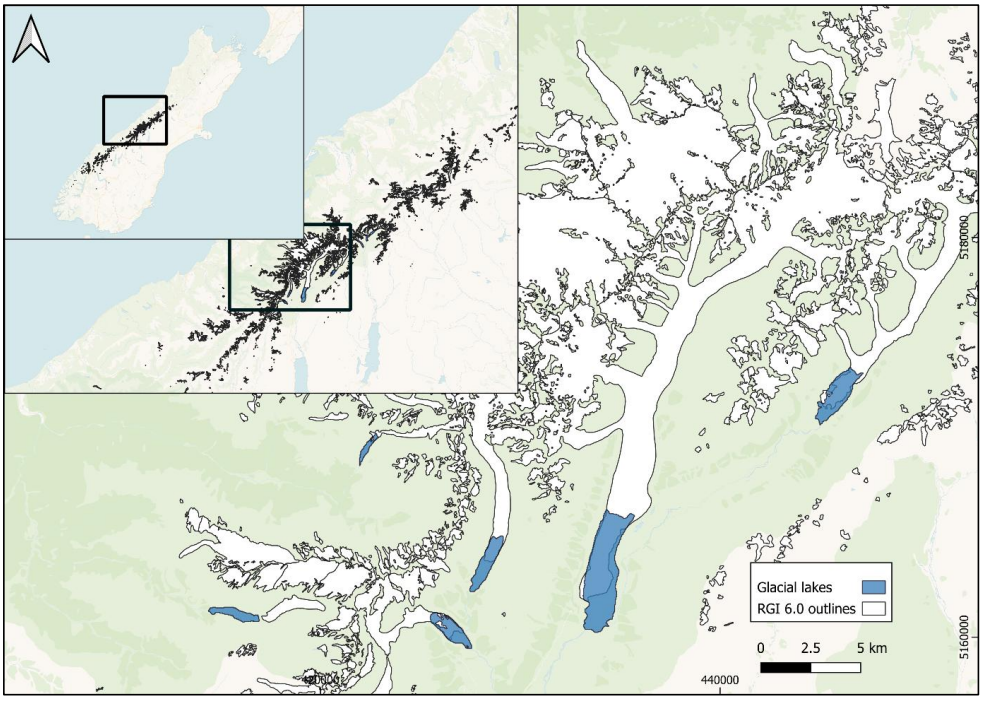


Figure 1 Glacial lakes in the study area for water-based indices comparisons. Basemap: CARTO (2026), Voyager, Map data © OpenStreetMap contributors.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

2 Data and methods

2.1 Data

The Landsat program (Landsat 1-9), spanning back to 1972, has the longest spatial and temporal record of multispectral satellite images, enabling the opportunity to document landcover changes for over five decades (Chen et al., 2017; Wulder et al., 2019, 2022). Launched through a collaborative effort between the US Geological Survey (USGS) and NASA, the Landsat program provides free data for assessing long-term changes to the Earth's environment (Markham et al., 2004; Masek et al., 2020). We use Landsat 4-9 Collection 2 Tier 1 calibrated top-of-atmosphere (TOA) reflectance images. Tier 1 images are used in this study because they are images of the highest available quality and are suitable for timeseries processing (Crawford et al., 2023; Earth Engine Data Catalog | Google for Developers, 2025). A Landsat 9 TOA image acquired on 2023/02/24 was selected as the training and testing image for segmentation and classification. The six feature classes selected for the training points are water, ice, dense vegetation, sparse vegetation, bare surface and shadow. To ensure a consistent application to all Landsat sensors, Landsat 4-7 bands are renamed to match Landsat 8 and 9 bands for B2-B7 and B10 (Table 1). These spectral bands were selected for segmentation and classification as they capture the key reflectance characteristics to delineate water bodies from surrounding snow, ice, vegetation and shadowed terrain. The Landsat 8-9 OLI/OLI-2 bands differ from the Landsat 4-7 TM/ETM+ bands through refined wavelength ranges in the Near Infrared, Short-Wave Infrared and Thermal Infrared bands. This reduced atmospheric interference, improved radiometric sensitivity and enhanced surface temperature retrieval (Irons et al., 2012). We use the NZ 8m Digital Elevation Model (DEM; LINZ Data Service, 2012), combined with the solar azimuth and elevation angles read from the Landsat metadata, to calculate a hillshade image to aid in the classification of different land cover.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Table 1 Renaming of Landsat 4, 5 and 7 bands in order (band names) to match the band names for Landsat 8 and 9.

Band	Landsat 4-5 Thematic Mapper (TM) and wavelength (µm)	Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and wavelength (µm)	Landsat 8-9 Operational Imager (OLI/-2) and wavelength (µm)	Renamed bands
Blue	B1 (0.45-0.52)	B1 (0.45-0.52)	B2 (0.45-0.51)	B2
Green	B2 (0.52-0.61)	B2 (0.52-0.61)	B3 (0.52-0.59)	B3
Red	B3 (0.63-0.69)	B3 (0.63-0.69)	B4 (0.64-0.67)	B4
Near Infrared	B4 (0.76–0.90)	B4 (0.76–0.90)	B5 (0.85-0.88)	B5
Shortwave Infrared 1	B5 (1.55-1.75)	B5 (1.55-1.75)	B6 (1.57-1.65)	B6
Shortwave Infrared 2	B7 (2.08-2.35)	B7 (2.08-2.35)	B7 (2.11-2.29)	B7
Thermal Infrared	B6 (10.40-12.50)	B6 VCID 1 (10.40-12.50)	B10 (10.6-11.19)	B10

2.2 Default method of OBIA lake extraction

The segmentation algorithm used in this study is Simple Non-Iterative Clustering (SNIC; Achanta and Susstrunk, 2017), derived from Simple Linear Iterative Clustering (SLIC; Csillik, 2017). SNIC segmentation involves simplifying an image into small clusters or connected pixels, known as objects or super pixels, that share common spectral characteristics (Achanta and Susstrunk, 2017). This approach has been widely used in land cover classification studies, with evidence documenting its advantages of efficiency and high accuracy (e.g., Tassi and Vizzari, 2020; Tu et al., 2020; Yang et al., 2021).

The SNIC implementation in GEE utilises different parameters which determine how an image is simplified into objects. These parameters include the seed size, compactness, connectivity, and *neighborhoodSize*. The seed size aids in determining the initial location or spacing of the cluster centres (Achanta and Susstrunk, 2017). The compactness determines how compact the pixels are with larger values increasing the compactness between pixels. The connectivity refers to the criteria to determine how neighbouring pixels are connected or considered part of the same object. We set the compactness to 0 to disable spatial distance weighting between pixels. We experiment using a connectivity of 4 or 8, considering only immediate neighbours, where a connectivity of 4 considers pixels connected if they share a horizontal or vertical edge while a connectivity of 8 considers pixels connected if they share an edge or a corner (Prathusha and Jyothi, 2018). The *neighborhoodSize* is used to investigate the effect of changing this parameter on boundary artefacts (Tassi and Vizzari, 2020).

RF classification was chosen as it is one of the most robust, computationally efficient and stable machine learning algorithms that has consistently performed well in land cover classification and for the detection of glacial lakes

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

in remote sensing imagery in GEE, and is suited to processing large, multi-decadal Landsat archives (Amani et al., 2020; Chen et al., 2017; Qayyum et al., 2020; Veh et al., 2018; Wangchuk and Bolch, 2020). Training data for the RF classifier was generated by sampling the object-based predictor bands produced by the SNIC segmentation at the location of manually interpreted points. Using these manually interpreted training points, they are divided into a 70/30% split where 70% of the training points are assigned to training and 30% are retained for independent testing. This approach ensures that the classifier learns from a representative subset while maintaining a separate dataset for accuracy assessment. Apart from changing the number of trees for experiments, we retain the default parameters of the RF in GEE (Table 2, Gorelick et al., 2017). Some studies have highlighted that more decision trees do not necessarily increase the performance of the classifier (e.g., Oshiro et al., 2012).

Table 2 Default Random Forest parameters in Google Earth Engine.

Argument	Type	Default parameter	Details
numberOfTrees	Integer	Must be specified	The number of decision trees to create.
variablesPerSplit	Integer	Null	The number of variables per split. If unspecified, uses the square root of the number of variables.
minLeafPopulation	Integer	1	Only create nodes whose training set contains at least these many points.
bagFraction	Float	0.5	The fraction of input to bag per tree.
maxNodes	Integer	Null	The maximum number of leaf nodes in each tree. If unspecified, defaults to no limit.
seed	Integer	0	The randomization seed.

2.3 Classification and segmentation experiments

To determine the impact of changing parameters on the resulting classification, we undertook a total of 48 experiments. The first 40 experiments were used to determine the most accurate combination of RF classifier parameters and the number of points per class by varying the number of training points and the number of trees classifier parameters (number of training points between 50-400, increasing by 50 each time, using either 10, 50, 100, 500, or 1000 trees. To analyse the optimal combination, we used a random sample of 100 manually classified points per class, taking into consideration the percentage accuracy (%, PA), user accuracy (%, UA) and $\hat{\kappa}$ statistic for each land cover class, with greater emphasis placed on the results for water. The percentage accuracy calculates the percentage proportion that was correctly classified in the classified OBIA map whereas the user accuracy refers to the probability of how likely that a given class or feature is correct. The $\hat{\kappa}$ statistic differs to the percentage and user accuracy as it corrects the accuracy of the classifier for change agreement, adjusting for agreement that could have happened by chance.

These experiments were applied to only the spectral information from the image (OLI bands 2-7, 10;**Table 1** -). Once the optimal classifier parameters were determined, we undertook an additional 8 experiments using the

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

optimal combination of classifier parameters and changed the segmentation parameters to identify the best combination of connectivity and *neighborhoodSize*. In this study, we refer to the OBIA classification across all feature classes as the multi-class classification and the OBIA classification of water against all other feature classes as the binary classification.

2.4 The effect of using input features on the OBIA classification

After determining the optimal combination of segmentation and classifier parameters using the original Landsat bands, we added additional input features to investigate how this changed the classification result. Due to the similar reflectance values of water and shadow, we used the NZ 8m Digital Elevation Model (DEM; LINZ Data Service, 2012), combined with the solar azimuth and elevation angles read from the Landsat metadata, to calculate a hillshade image. Both the hillshade and DEM slope were then added as input features for the classifier. Both the MNDWI and NDWI were also added as features, along with the Normalised Difference Vegetation Index (NDVI), a widely used index that utilises the relationship between chlorophyll reflectance in the infrared and red spectral bands (Myneni et al., 1995; Pettorelli et al., 2005). The export of band reflectance values in GEE is used here to aid in the investigation of how the influence and importance of bands (B2-B7 and B10) and input features (MNDWI, NDVI) changes through Landsat 4-9 OBIA classifiers. Feature importance analysis helps determine which bands and input features are the most influential when running the OBIA classifier. The exported band and input feature values identified by the mean, standard deviation and median values of the training points for every class.

2.5 Applying the OBIA classification to other Landsat satellites

As a final comparison experiment, we compared the result of applying a classifier trained on a single Landsat 9 OLI-2 image against the result of applying a classifier trained on a set of five images, one from each Landsat sensor (Table 3, TM4, TM5, ETM+, OLI, OLI-2). This experiment is intended to determine whether a classifier trained on a single image can accurately classify the images from other sensors, or whether it is necessary to use examples from each sensor in training. These experiments aid in determining whether increasing the amount of training images across the collection improves the accuracy of the classifier and the detection of water across the Landsat era.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Table 3 Table highlighting the Landsat ID, path and row and CLOUD_COVER according to GEE properties for individual Landsat satellite images used in this study.

Landsat satellite	Landsat ID	Path / Row	Cloud Cover (%)
LT04	LT04_L1TP_075090_19901230_20200915_02_T1	75 / 90	9
LT05	LT05_L1TP_075090_20070204_20200831_02_T1	75 / 90	2
LE07	LE07_L1TP_075090_20030201_20200916_02_T1	75 / 90	5
LC08	LC08_L1TP_075090_20220213_20220222_02_T1	75 / 90	5
LC09	LC09_L1TP_075090_20230224_20230308_02_T1	75 / 90	4.6

2.6 Calculating lake area outlines and accuracy assessment

QGIS is a free and open-source cross platform desktop geographic information system application that supports viewing, editing, printing and analysis of geospatial data (QGIS geographic information system). The OBIA maps for all 5 images underwent post-processing steps in QGIS, using the WGS 84 UTM zone 59S coordinate reference system. For calculating the area of each glacial lake, all vector objects present in the lake, excluding land, were merged. We used the Semi-Automatic Classification Plugin (SCP) in QGIS (Congedo, 2013, 2021) for accuracy assessment and calculating lake area uncertainty and applied this to every OBIA map created for every image in this study. The accuracy for the automatically generated feature class outlines created by the OBIA was evaluated by generating 100 random points within each feature class and then comparing this to a manual classification for each point. The classification standard error is calculated by the tool following the approach detailed by Olofsson *et al.*, (2014). We report this here as a 95% confidence interval.

2.7 Comparison of water-based indices with OBIA results

We compared approaches using water-based indices (NDWI, MNDWI, AWEI_sh) and a RF PBIA approach against the results of the final OBIA classification for 5 Landsat images and applied this to six glacial lakes in the study region (Figure 1). We select one image per Landsat sensor to compare the results of our OBIA across the Landsat archive and ensured that the selected images were restricted to the local ablation season with cloud cover below 10% (December-March). For each index, we calculate the mean value and standard deviation (±) within the individual lake boundaries. The threshold value for each index is chosen based on the lowest threshold found for the glacial lakes (Table 4). This is then applied to the whole image collection without the need of manually changing the threshold every time to make it a fair comparison with the OBIA classification. Therefore, thresholds for the indices are larger than what are applied for other studies that may only apply the threshold to one individual lake or sensor. Additionally, the comparison is made using different glacial lakes in the study area, which have vastly different thresholds due to different spectral properties caused by differences in lake turbidity. For the RF PBIA classification, we used the same six feature classes and the same 250 points per class for a more direct comparison. Finally, we manually delineated the six glacial lakes in the study area as a reference for the "true" area of the glacial lakes, applying a buffer of ±15 m to estimate the uncertainty in the manual delineation.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Table 4 The mean value of the index and standard deviation (±) of individual glacial lakes within the study area, and the threshold value selected for each index derived from the lowest value from all glacial lakes.

Index	Tasman	Murchison	Mueller	Hooker	Strauchon	Douglas South	Threshold
NDWI	0.57 ± 0.1	0.58 ± 0.11	0.4 ± 0.1	0.5 ± 0.12	0.62 ± 0.33	0.55 ± 0.17	0.4 >
MNDWI	0.96 ± 0.15	0.94 ± 0.21	0.9 ± 0.24	0.94 ± 0.22	0.79 ± 0.38	0.89 ± 0.28	0.79 >
AWEI_sh	0.67 ± 0.15	0.65 ± 0.19	0.39 ± 0.16	0.58 ± 0.2	0.48 ± 0.42	0.45 ± 0.22	0.39 >

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

3 Results

3.1 Classification parameters

Figure 2 shows heatmaps of the overall accuracy (%) and $\hat{\kappa}$ statistic for the multi-class OBIA classification for the training and testing image. When comparing the number of trees with the number of training points per class in both the overall accuracy and $\hat{\kappa}$ statistic heatmap, the optimal amount of training points per class is identified as 250, with four of the five highest overall accuracies residing in this block. This is supported by the heatmaps of overall accuracy for the binary OBIA classification $\hat{\kappa}$ statistic where 250 training points per class remains the optimum amount of training points with three out of the five highest scores for both, respectively. The heatmap the percentage accuracy (%) of water shows that there is not an optimum number of training points, with between 150-400 points showing high accuracies for water detection (Figure 3). However, the heatmap of user accuracy and $\hat{\kappa}$ statistic of water shows a distinct pattern where three out of the five highest scores are documented for 250 points. We find that the optimum amount of training points needed per class and specifically water is 250 points. Identifying the optimum number of trees combined with the 250 training points per class is a challenging task due to similarly high values across the land classes when comparing the overall accuracy and $\hat{\kappa}$ statistic of both the classifier and water.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

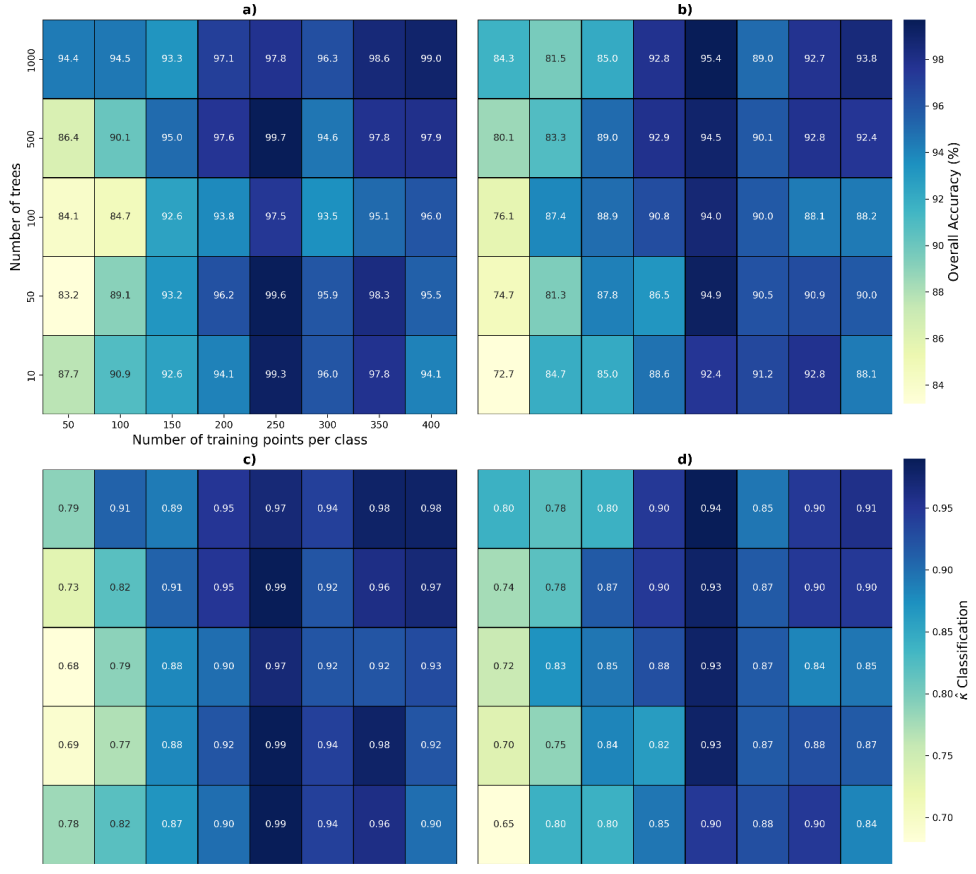


Figure 2 Heatmaps illustrating comparing the changes in overall accuracy (%) and $\hat{\kappa}$ statistic when changing the number of trees and the number of training points per class showing for the training and testing image. A) overall accuracy for the multi-class OBIA classification; B) the overall OBIA classification of water detection compared to the other 5 classes, C) the $\hat{\kappa}$ statistic for the overall multi-class OBIA classification and D) the $\hat{\kappa}$ statistic for water detection compared to the other 5 classes.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

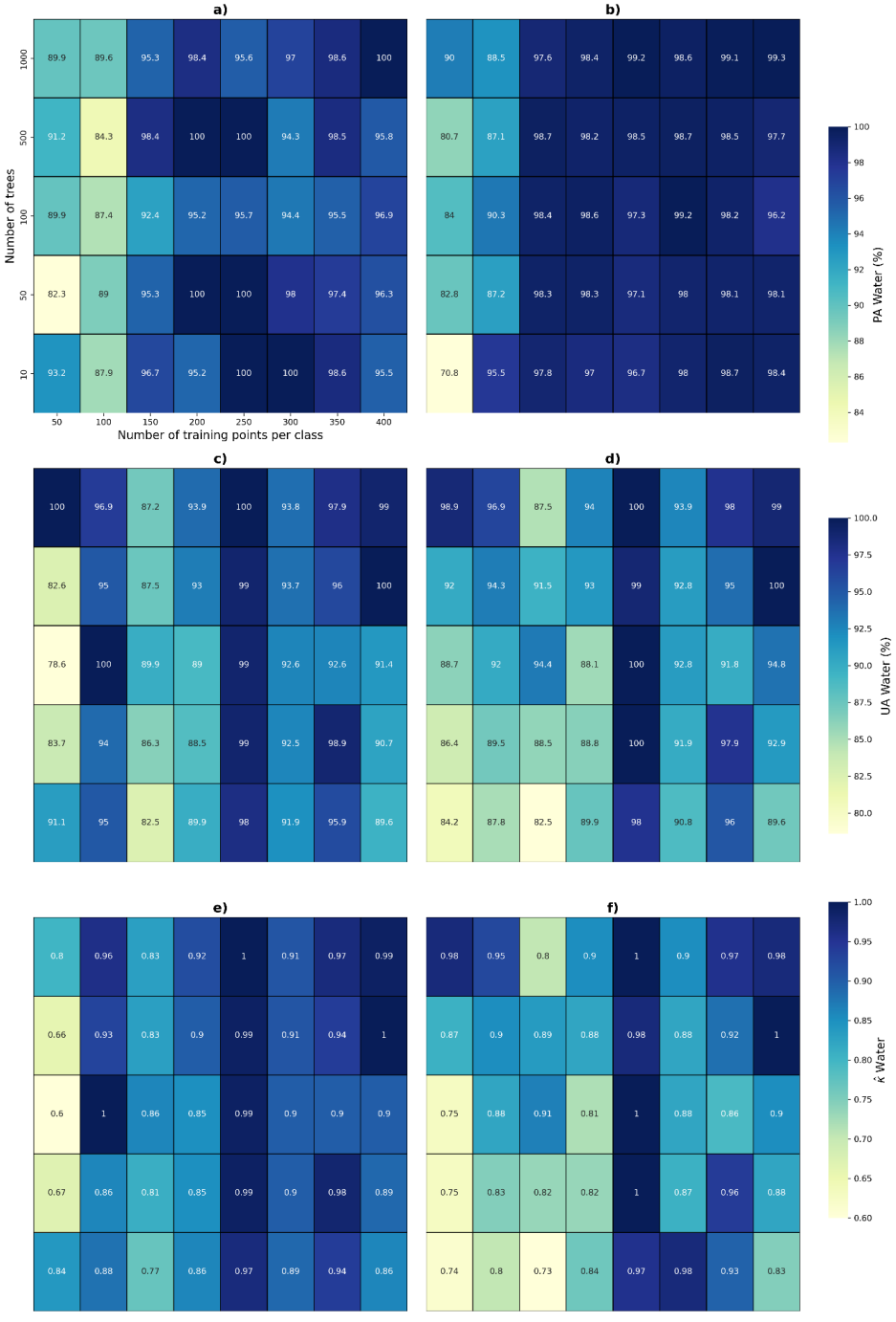


Figure 3 Heatmaps illustrating heatmaps of the (a) percentage accuracy of water in the OBIA classification (%), (b) percentage accuracy of water against all the other classes combined (c) the user accuracy of water OBIA classification (%), (d) the user accuracy of water against all the other classes combined (e) $\hat{\kappa}$ statistic of water in the OBIA classification and (f) the $\hat{\kappa}$ statistic of the binary OBIA classification.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Given the similarity in performance for these combinations, we also investigated the differences in Batch Computer Usage as a function of the number of classifier trees using the Earth Engine Compute Unit (EECU) in seconds, taking 10 measurements each and calculating the mean and standard deviation. The EECU is a mechanism for representing an amount of instantaneous processing power. The classifier with 10 trees has the lowest batch computer usage of 13464.2 ± 1165.8 EECU-seconds, while 1000 trees require three times as many resources, with 25167 ± 1624.9 EECU-seconds (Table 5). Given the similarity in both classifier performance and resource use, we conclude that increasing the number of trees from 100 to 500 does not provide sufficient improvement to warrant the additional resource usage.

Table 5 Table showing the difference in the average run time runtime (minutes) and batch computer usage (EECU-seconds) between different number of trees.

Number of trees	Runtime (minutes)	Batch computer usage (EECU-seconds)
10	8.48 ± 1.48	13464.2 ± 1165.8
100	9 ± 2.12	13805.4 ± 873.5
500	9.1 ± 2.12	17030.3 ± 2284.4
1000	11.4 ± 1.42	25167 ± 1624.9

3.2 Segmentation parameters

Table 6 shows the result of the segmentation experiments of the OBIA created in GEE. When comparing the combination of segmentation parameters for the OBIA classification result for all multi-class, and binary classifications, the combination of a connectivity of 4 and a *neighborhoodSize* of 128 excelled in all accuracy measurements. This is highlighted by achieving the highest accuracies of overall accuracy (%), percentage and user accuracy of water for both the multi-class OBIA and binary classification of 94%, 97.4%, 100% and 97.5%, 99.3 and 96.8, respectively (Table 6). Our results show that changing the connectivity from 4-8 is more detrimental to the accuracy measurements than changing the *neighborhoodSize*, which is highlighted by the large disparity between the combinations of 4-128 and 8-128 in comparison to the increase of *neighborhoodSize* from 128 to 256. Due the superiority of the combination of 4-128 for all measurements of accuracy we find that the optimal combination of segmentation parameters for the OBIA classification is a connectivity of 4 and a *neighborhoodSize* of 128.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Table 6 Table showing the comparison of changing the segmentation parameters (connectivity and neighborhoodSize) paired with 100 trees to identify the optimum combination to yield the highest overall accuracy (%) and κ̂ statistic in the multi-class and binary OBIA Classification using the original bands and input features on the training image.

Connectivity and Neighborhood Size	Overall accuracy (%)	K̂ statistic	Percentage Accuracy Water (%)	User Accuracy Water (%)	K̂ statistic Water
Multi-class OBIA classification					
4-128	94	0.93	97.4	100	1
4-256	86.3	0.83	96.1	95	0.93
8-128	79.8	0.73	96.1	78	0.79
8-256	83.8	0.79	97.1	85.6	0.88
Binary classification					
4-128	97.5	0.97	99.3	96.8	0.95
4-256	96.9	0.95	96.4	97.8	0.94
8-128	86.9	0.82	87.1	93.9	0.89
8-256	90.9	0.87	90.6	95.7	0.93

3.3 Impact of input features

Our study uses the combination of 250 feature points per class with 100 trees to produce the OBIA classifier. The importance of adding the input features (hillshade, NDVI, MNDWI, slope) is highlighted through the improvement of the classification of land types in comparison with the original image bands alone. This is displayed in Figure 4 where much less water is misclassified in the classifier when the input features are added against when we use only the original bands (B2-B7, B10). Additionally, the comparison of the overall accuracy, κ̂ statistic percentage and user accuracy of the multi-class classification and the binary classification show the improved results of adding these input features. This is evident when comparing an overall accuracy and a κ̂ statistic of 95.7% and 0.94 for the multi-class classification and 98.7% and 0.98 for the binary classification when adding input features in comparison to 91.1% and 0.88 and 96.7% and 0.96 without the input features (Table 7). The input features improve the overall accuracy of the classifier and improves the water detection in the original training image and is used in this study. When comparing the misclassified pixels before and after the inclusion of input features, we find that the inclusion of input features reduced the number of misclassified pixels by 50.7% (Table 8). Before the inclusion of the input features, the highest misclassified class for water was shadow with 26 pixels, representing 38.8 % of the total misclassified pixels. After the inclusion of input features, no shadow pixels were misclassified as water, indicating the improved differentiation between these two classes with the input features.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

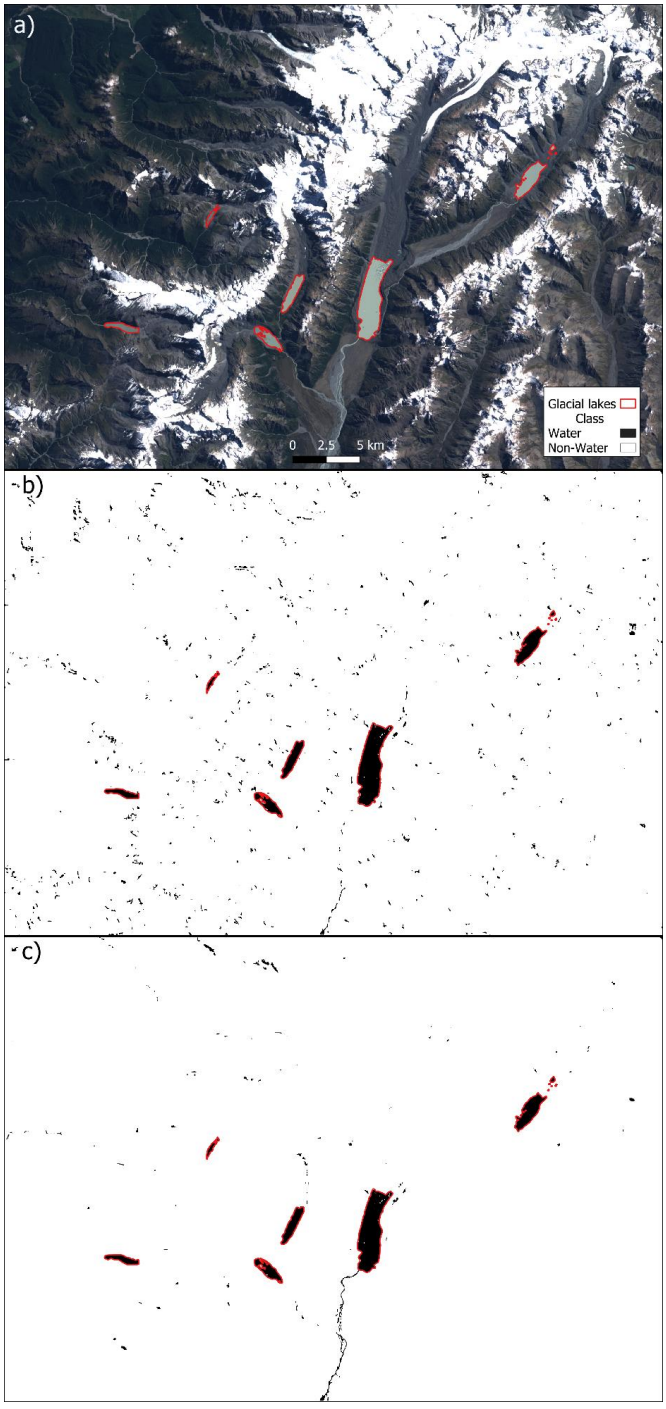
Table 7 Overall accuracy (%), κ̂ statistic the percentage accuracy water (PA, %), user accuracy (UA, %) and K̂ water for the multi-class OBIA classification and the binary classification for the original bands against the input features for the study area of Figure 2.

Input features	Overall accuracy (%)	K̂ statistic	Percentage Accuracy Water (%)	User Accuracy Water (%)	Kappa statistic Water
Multi-class classification					
No	91.1	0.88	97.6	92.9	0.88
Yes	95.7	0.94	99.2	100	1
Binary classification					
No	96.7	0.96	94.5	93.2	0.94
Yes	98.7	0.98	100	98	0.97

Table 8 The number and class of misclassified pixels per class for the B (original bands) and A (input features) when running accuracy assessments on the OBIA classifier.

Class	Misclassified land cover											
	Water		Ice		Dense vegetation		Sparse vegetation		Bare Surface		Shadow	
	B	A	B	A	B	A	B	A	B	A	B	A
Water	0	0	0	0	0	0	0	0	1	2	26	0
Ice	1	0	0	0	0	0	0	0	1	2	2	5
Dense vegetation	1	0	0	0	0	0	4	4	5	2	0	2
Sparse vegetation	0	0	0	0	6	2	0	0	3	2	0	0
Bare Surface	0	0	0	0	3	1	1	1	0	0	5	2
Shadow	5	1	0	1	2	5	0	0	2	3	0	0
Total	7	1	0	1	11	8	5	5	12	11	33	9

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026



340

341 Figure 4 a) The study area image (Figure 1) and the comparison of the binary OBIA classification map b) before
342 and c) after applying additional input features. Landsat 9 Collection 2 Tier 1 TOA imagery courtesy of the U.S.
343 Geological Survey.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

3.4 Application of the OBIA classifier to all Landsat sensors

Table 9 shows the comparison of the overall accuracy (%), κ̂ statistic the percentage accuracy water (PA, %), user accuracy (UA, %) and κ̂ water for the multi-class and binary OBIA classification when applied to Landsat images acquired using different sensors. Our results highlight that the OBIA classifier trained and evaluated on the Landsat 9 image acquired on 24/02/2023 can be applied on any image in the collection from any Landsat sensor. The overall accuracy of the multi-class OBIA classification for each image is higher than 93%, indicating that the OBIA classifier was successfully able to classify different land types accurately. This is further supported by values of 0.9 exceeded for all Landsat images which indicates that our results are in very good agreement that the multi-class OBIA classifier outputs reflect the true feature classes. Additionally, exceptionally high PA and UA and K̂ for water supports our results that our OBIA classifier can be applied to any image across the Landsat era in the Southern Alps.

Table 9 The comparison of the overall accuracy (%), κ̂ statistic the percentage accuracy water (PA, %), user accuracy (UA, %) and K̂ Water for the OBIA classifier.

Satellite ID	Overall accuracy (%)	κ̂ statistic	PA Water (%)	UA Water (%)	K̂ Water
Multi-class classification					
LC09-2023-02-24	95.7	0.94	99.2	100	0.94
LC08-2022-02-13	93.2	0.92	93.2	93.2	0.92
LT05-2007-02-04	97.3	0.96	97.3	97.3	0.96
LE07-2003-02-01	93.7	0.91	93.7	93.7	0.91
LT04-1990-12-30	93.3	0.91	93.3	93.3	0.91
Binary classification					
LC09-2023-02-24	98.7	0.98	100	98	0.97
LC08-2022-02-13	98.3	0.98	98.9	98.1	0.98
LT05-2007-02-04	98.7	0.98	100	98.6	0.98
LE07-2003-02-01	99.1	0.99	98.9	99.8	0.99
LT04-1990-12-30	98.4	0.98	98.7	98.9	0.98

3.5 The effect of training multiple images to the OBIA classifier

Table 10 shows the evaluation metrics for the classifier that was trained on five images. These results show that the OBIA classifiers trained and merged on five images produced weakened results for all measurements when compared to the singularly trained OBIA classifiers (Table 10, Table 11). This is supported by Landsat 4-7 images which have significantly lower overall accuracy and κ̂ statistic in comparison to the training of the OBIA classifier on a single image. This comparison highlights the superiority of training the OBIA classifier on one image compared to five images. This shows that the five image OBIA classifier across the Landsat collection

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

misclassifies glacial lakes as other land types and therefore the application of a single image trained on the OBIA classifier is superior to apply to this study across the whole Landsat era (Figure 5).

Table 10 The comparison of the overall accuracy (%), κ̂ statistic the percentage accuracy water (PA, %), user accuracy (UA, %) and K̂ statistic for water for the individually trained OBIA classifiers per Landsat sensor.

Satellite ID	Overall accuracy (%)	κ̂ statistic	PA Water (%)	UA Water (%)	K̂ Water
Multi-class OBIA classification					
LC09-2023-02-24	95.7	0.94	99.2	100	1
LC08-2022-02-13	94.3	0.92	98.2	98.7	0.98
LT05-2007-02-04	95.9	0.94	100	100	1
LE07-2003-02-01	96.7	0.95	100	98.9	0.98
LT04-1990-12-30	94.6	0.93	99.1	98.0	0.97
Binary OBIA classification					
LC09-2023-02-24	98.7	0.98	100	98	0.97
LC08-2022-02-13	98.5	0.97	100	98.9	0.98
LT05-2007-02-04	98.3	0.97	98.9	97.0	0.96
LE07-2003-02-01	98.4	0.98	100	98	0.98
LT04-1990-12-30	98.2	0.97	100	95.8	0.93

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Table 11 The comparison of the overall accuracy (%), κ̂ statistic the percentage accuracy water (PA, %), user accuracy (UA, %) and K̂ statistic for water for the multiple images trained to a singular OBIA classifier.

Satellite ID	Overall accuracy (%)	κ̂ statistic	PA Water (%)	UA Water (%)	K̂ Water
Multi-class OBIA classification					
LC09-2023-02-24	92.5	0.9	96.8	97.8	0.96
LC08-2022-02-13	91.5	0.88	99.7	97.8	0.96
LT05-2007-02-04	88.6	0.85	87.3	94.3	0.91
LE07-2003-02-01	87.5	0.84	86.4	94.4	0.9
LT04-1990-12-30	84.9	0.8	84.4	92.2	0.9
Binary OBIA classification					
LC09-2023-02-24	98.5	0.98	100	99	0.98
LC08-2022-02-13	98.5	0.98	98.2	100	1
LT05-2007-02-04	98.9	0.98	96.9	99.0	0.99
LE07-2003-02-01	97.8	0.97	96.9	99.0	0.98
LT04-1990-12-30	96.6	0.95	95.6	95.9	0.95

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

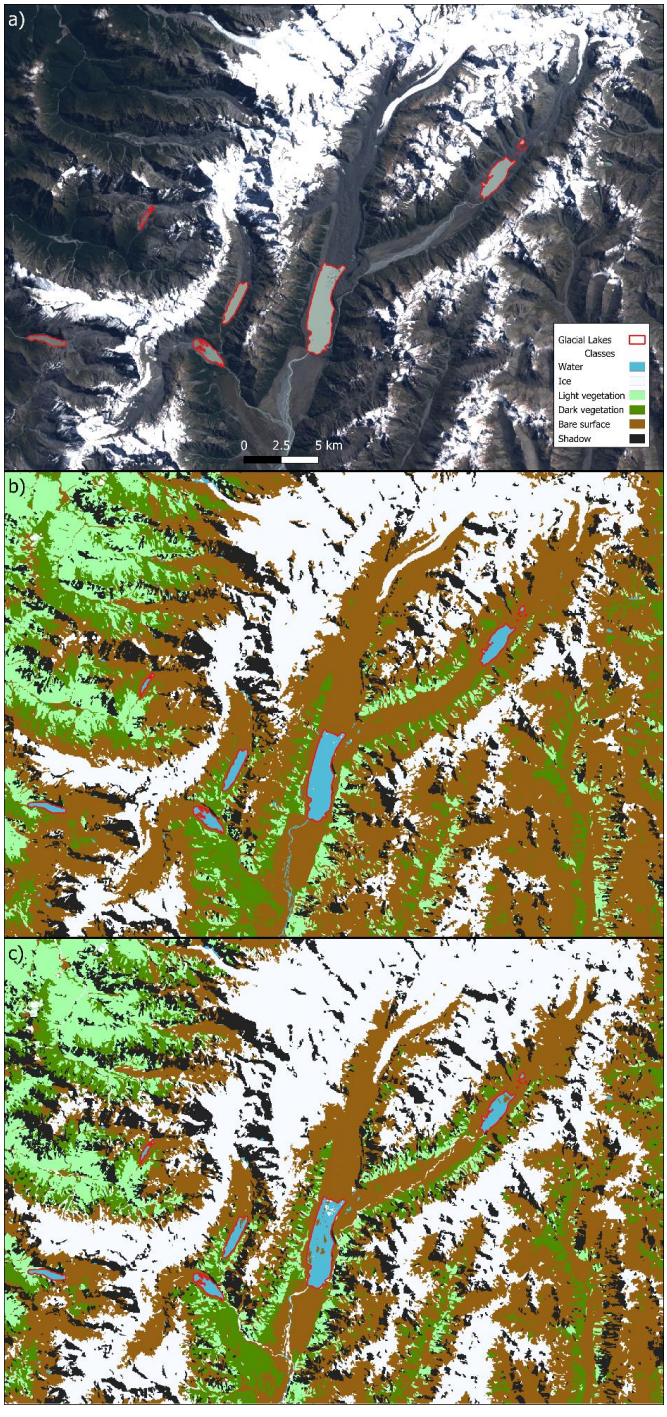


Figure 5 a) The original training image, b) the multi-class OBIA classification trained on the original training image and c) The multi-class OBIA classification trained on five images. Landsat 9 Collection 2 Tier 1 TOA imagery courtesy of the U.S. Geological Survey.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

3.6 Water extraction indices comparison with the OBIA

Table 12 shows the comparison of other water extraction indices with the binary OBIA classification. We identify the superiority of the OBIA which has an overall accuracy of 95.4 % for the binary OBIA classification in comparison with 67.8 % for the NDWI, 54.8 % for the MNDWI and 56.2% for the AWEI_sh. Additionally, results shown in Table 12 indicate that the OBIA is superior in identifying water, as the overall accuracy was 98.9 %, in comparison of 28.6 % for the NDWI, 5.1 % for the MNDWI and 7.4 % for the AWEI_sh. The standard error for water using the OBIA is significantly less than the other water extraction indices, with an average standard error of 1.3 % in comparison to 3.2 %, 3.5 % and 3.4 % respectively, for the NDWI, MNDWI and AWEI_sh.

Table 13 shows the comparison of the results derived by applying an OBIA and a PBIA. The results show the superiority of the OBIA which has an average of 97.9 % in comparison to 74.5 % for the PBIA. Additionally, our results indicate that an OBIA method produces higher overall producer's accuracies of 98.7 %, in comparison to 59 % of the PBIA. The standard error for water using the OBIA is significantly less than the PBIA, with an average standard error of 0.57 % compared to 1.63 % respectively, for the PBIA method. Therefore, we can conclude that the OBIA method is able to delineate water bodies at higher accuracies throughout the whole Landsat collection in comparison to a PBIA method.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Table 12 The results of overall accuracy (OA), producer's accuracy of water (PA) and standard error (SE) for the water class as a % for different classification methods across multiple Landsat images.

Date	Satellite image	Method											
		NDWI			MNDWI			AWEI_sh			OBIA		
		OA	PA	SE	OA	PA	SE	OA	PA	SE	OA	PA	SE
24/02/2023	LC09-2023-02-24	79	54.3	2.76	53.5	5	3.5	58.3	11.2	3.43	95.7	99.2	1.2
13/02/2022	LC08-2022-02-13	70.5	33	3.10	56	7.4	3.47	58.5	11.7	3.42	94.3	99.3	1.25
04/02/2007	LT05-2007-02-04	65.5	20.9	3.30	59	3.5	3.47	55	4.3	3.49	95.9	98	1.4
01/02/2003	LE07-2003-02-01	60	14	3.42	53.5	3.1	3.51	53	4	3.45	96.7	100	1.12
10/01/2001	LT04-1990-12-30	64	21	3.4	52.1	6.4	3.51	56	6	3.43	94.6	98	1.4
Average		67.8	28.6	3.2	54.8	5.1	3.5	56.2	7.4	3.4	95.4	98.9	1.3

Table 13 The results of overall accuracy (OA), producer's accuracy of water (PA) and standard error of water (SE) as a % of water extraction indices (NDWI, MNDWI, AWEI_sh) for the binary OBIA classification using the 5 images across the Landsat collection.

Date	Satellite image	Method					
		OBIA			PBIA		
		OA	PA	SE (%)	OA	PA	SE (%)
24/02/2023	LC09-2023-02-24	98.7	97.4	0.38	80	67	1.28
13/02/2022	LC08-2022-02-13	96.2	98.7	0.49	88.4	73	0.9
04/02/2007	LT05-2007-02-04	98.5	99	0.65	73.1	55	1.91
01/02/2003	LE07-2003-02-01	98.1	100	0.69	63.6	48	2.11
10/01/2001	LT04-1990-12-30	98.2	98.2	0.68	67.2	52	1.97
Average		98.7	97.4	0.38	80	67	1.28

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

4 Discussion

4.1 Optimal classification and segmentation parameters for glacial lake detection

We find that 250 points per feature class and 100 random forest decision trees produced the best results for overall accuracy of the OBIA classifier. This is supported by other studies which highlighted that more decision trees does not necessarily increase the performance of the classifier, and small gains in accuracy do not compensate for longer computational processing time (Oshiro et al., 2012; Probst and Boulesteix, 2018). The number of training samples per class is one of the most important factors that can affect the results of an accuracy assessment, alongside the number of classes in the classification, the ability of the training data to characterise the feature classes mapped, and the dimensionality of the data (Congalton, 1991; Foody and Mathur, 2004; Millard and Richardson, 2015; Pal and Mather, 2003). We justify the use of 250 points per feature class due to the area of our training image, and the number of training points is adequate and representative of the entire landscape ensuring that all feature classes were weighted equally in the classification. It is important to keep all training points for feature classes equally weighted because machine learning algorithms such as the RF may favour classes with more training data compared to classes with less training data, leading to an over or under-representation of feature classes in the resultant map (Millard and Richardson, 2015).

Our results indicate that the RF classifier is sensitive to changes in the number of training samples, highlighted by abrupt changes in overall accuracy of the OBIA classifier for all the multi-class and binary OBIA classification (Figure 2, Figure 3). This is supported by Millard and Richardson (2015) who found that the RF is sensitive to the size of the training data set and that the training data set must be randomly and spatially distributed across the whole image to represent the actual class proportions in the landscape. We suggest that the number of 250 feature points is not the standard amount needed for the classification of an OBIA classifier but is the optimal representative number of points per class for our image and feature classes. We identify that the number of points below or above this threshold may be under or over-representing certain feature classes over others and future studies should take caution in the amount of training samples needed per feature class, as this is dependent on the amount of feature classes identified and the landscape on the image.

The optimal segmentation parameters for this study were the combination of a connectivity of 4 and a *neighborhoodSize* of 128. Previous studies have used a connectivity of 8 for both lake and land cover mapping for OBIA analysis in GEE (Karakuş, 2024; Shafizadeh-Moghadam et al., 2021). However, we found that a connectivity of 4 paired with 128 *neighborhoodSize* significantly improves the overall accuracy of the OBIA classifier in comparison to a connectivity of 8 with 128 *neighborhoodSize* (Table 6). Studies have highlighted that there is no correct combination because the SNIC outputs are highly variable depending on the visualisation scale (Tassi and Vizzari, 2020). Therefore, experiments should be run on different satellite sensors within GEE to identify the correct combination of segmentation parameters, particularly for SNIC segmentation parameters applied to Sentinel 2 (10m resolution) or future Landsat sensors with increased resolution.

4.2 The inclusion of input features in the OBIA classifier

We found that the inclusion of input features to the OBIA classifier improves the overall accuracy of the OBIA classifier and the detection of water in the image. This is highlighted with the decrease in pixel misclassification by 50.7% from 67 misclassified pixels without the inclusion of input features to 34 misclassified pixels when

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

adding input features and indices (Table 8). The addition of a hillshade layer and the MNDWI to the classifier eliminated the misclassification of shadow as water in the image. Studies that have used GEE to delineate glacier area changes have emphasised the need to include additional datasets such as DEMs in OBIA classifier codes to improve the efficiency of the resultant map (Ali et al., 2023). We addressed this need and showed that the inclusion of a DEM is important for delineating hillshade and slope in our OBIA classifier. It is important to differentiate shadow and water in an image, particularly when training a whole image that has both darker lakes that have similar low reflected spectral signals for the visible and infrared bands as shadows (Yue et al., 2023). Our study supports existing research (e.g. Weekley and Li, 2021; Chen and Zhao, 2022; Yue *et al.*, 2023) that a DEM must be included in any OBIA classifier to fully automate and delineate highly accurate lake mapping. This is shown by the vast improvement in the exclusion of the topographic shadows when adding a DEM that aids in calculating a hillshade and can be applied successfully to any given image in the whole Landsat collection (Figure 4).

We find increased accuracy in water detection in the OBIA classifier when adding input features (Table 7,Table 8). The inclusion of the MNDWI in the OBIA classifier is important for differentiating water and shadow and has been used to map water bodies in many regions with relatively stable thresholds in comparison with other water bodies (Dou et al., 2023; Feyisa et al., 2014; Ji et al., 2009; Singh et al., 2015; Zhang et al., 2015; Zhao et al., 2018). The MNDWI is added without thresholding, due to the sporadic differences in thresholds between Landsat sensors and spectral values between glacial lakes as well as the spectral value differences between the ocean and other lakes in the image. The NDVI is included in this study alongside the MNDWI to aid in the differentiation between water and vegetation. The combination of using the NDVI and MNDWI has been successful in differentiating water bodies from vegetation at high accuracies using a decision tree model (Gujrati and Jha, 2018) as well as calculating the variations in the area of surface water bodies in the Huai River Basin in China by using the NDVI, MNDWI and the enhanced vegetation index (EVI; Jiang et al., 2008; Xia et al., 2019). Our study goes a step further by incorporating these indices with input features in hillshade and slope and identifies the improvement of the overall accuracy of the OBIA classifier by producing additional information for the code to differentiate between land classes. We show that the combination of input features in an OBIA approach is vital in improving the overall classification of land classes in our training image (Figure 6). This is demonstrated by an improvement in the delineation of water in an image as well as ice, different vegetation, and shadow bodies at a high accuracy.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

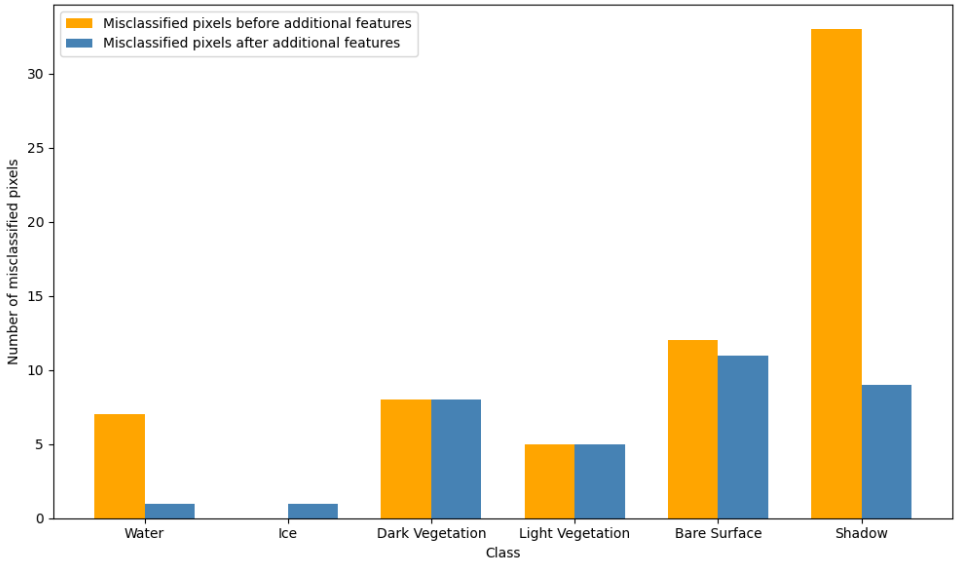


Figure 6 The comparison of the number of misclassified pixels before and after applying input features.

4.3 Comparison of training and evaluating one image against multiple for the OBIA Classifier

Our study identifies that the delineation of water bodies is superior when training and testing one image compared to training and testing 5 images from different Landsat sensors. To understand why this was the case, we investigated the similarities and differences between feature values for all the original bands and input features for the training points of water and other land classes. Previous studies identified the differences in radiometric resolution of Landsat sensors as a factor, calculating slightly higher accuracies for land classification studies using images with higher radiometric resolution compared to lower resolution data (Tucker, 1980). Recent studies have found that lower radiometric resolution data can be used safely in object-based classification (Rama Rao et al., 2007; Verde et al., 2018). Our results suggest that radiometric resolution is not a challenge for accurate classification. Radiometric sensitivity has improved between Landsat 4 which has an 8-bit dynamic range with 256 grey levels 0 – 255, compared to Landsat 8 and 9 with 12-bit dynamic range and 4096 grey levels (Wulder et al., 2019). Some research has shown the significant advancements made by Landsat 8 and 9 in radiometric sensitivity for the glaciological community, especially for mapping ice flow in glaciers and ice sheets as well as glacial lake bodies (Fahnestock et al., 2016; Kingslake et al., 2017; Wulder et al., 2019). Studies using Landsat 8 imagery with increased radiometric resolution identified improvements in the mapping of meltwater lakes on the ice sheets, by capturing finer variations in reflectance, which is important to differentiate water or dark surfaces and shadow (Fahnestock et al., 2016; Jeong and Howat, 2015; Mouginot et al., 2017; Wulder et al., 2019).

When combining five trained OBIA classifiers from five Landsat sensors into a singular OBIA code we find that the overall accuracy of the classifier decreases in comparison to training the five images individually (Table 10, Table 11). Here, we investigate the feature class values of individual bands and input features (MNDWI, NDVI) and find that the differences in the overall accuracy of the OBIA classification is due to the heterogeneous

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

differences in values that are skewing the overall classification, causing overlap and leading to misclassification (Figure 7). Zhai, Qu and Hao, (2018) highlight that the primary method of land cover classification obtains metrics such as the mean, median and standard deviation values of feature values for classification in a supervised approach. They find that any changes to the mean, median and standard deviation values of input features can have a significant impact on the classification of individual land classification and the overall OBIA classifier. We identify the same outcome when comparing all the Landsat sensors, but specifically Landsat 9 with all sensors for water classification (Figure 7).

The distribution peaks for individual bands are wider for all sensors compared to Landsat 9, which may be the cause of the misclassification of land classifications as there is more overlap between differing land classes. Wider peaks are associated with an increased range of values that are used to classify land cover, and multiple wide peaks that overlap increase the likelihood of misclassification through an increased number of land cover classes share a similar wide range of values. For example, when comparing the distribution peaks of the near-infrared band (NIR), Short-Wave Infrared 1 (SWIR 1) and Short-Wave Infrared 2 (SWIR 2; B5, B6 and B7) we find significantly less overlap between water and other land classes for Landsat 9 than all sensors (Figure 8). Studies have found that the NIR band has the largest difference in reflectance values between water and non-water bodies, as water absorbs more in this wavelength where non-water, specifically shadow, has a higher reflectance (Feyisa et al., 2014; Mondejar and Tongco, 2019). This is emphasised in our comparison of one image OBIA classification compared to five, with significantly more misclassification between water and other land classes in the five image OBIA classification (Figure 9). Therefore, the five image OBIA classifier for all sensors produces less accurate maps due to the increased density of NIR, SWIR 1 and 2 and other bands at differing values, creating widening peaks that overlap with other land classification and driving misclassification.

Another factor that contributes to a lower overall accuracy for the five image OBIA classifier is temporal changes in land cover. Our five trained and tested images for the OBIA are split across 34 years and are taken in the month of February, apart from the Landsat 4 image in December due to low availability of satellite images. Studies have suggested that seasonal variations and irregular intervals of satellite imagery captured by multispectral images may influence the overall OBIA classifier (Bueno et al., 2019). For example, for the Landsat 4 image in December we may see difference in individual land cover reflectance values, particularly for fresh snow vs partially melted snow and leaf-off vs leaf on conditions. This may produce wider distribution of values of each land class and leading to an increase in the misclassification between classing, causing overlap reflected in our study.

Training and evaluating one Landsat 9 image produces a rapid and efficient OBIA classifier that can be applied to the whole collection. Due to the complexities of radiometric sensitivity, heterogenous values for bands and indices and temporal changes in land cover throughout the study, training and testing five images across the collection returns lower accuracies for the OBIA classifier and leads to increased misclassification between land classes and most importantly in water detection in the image.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

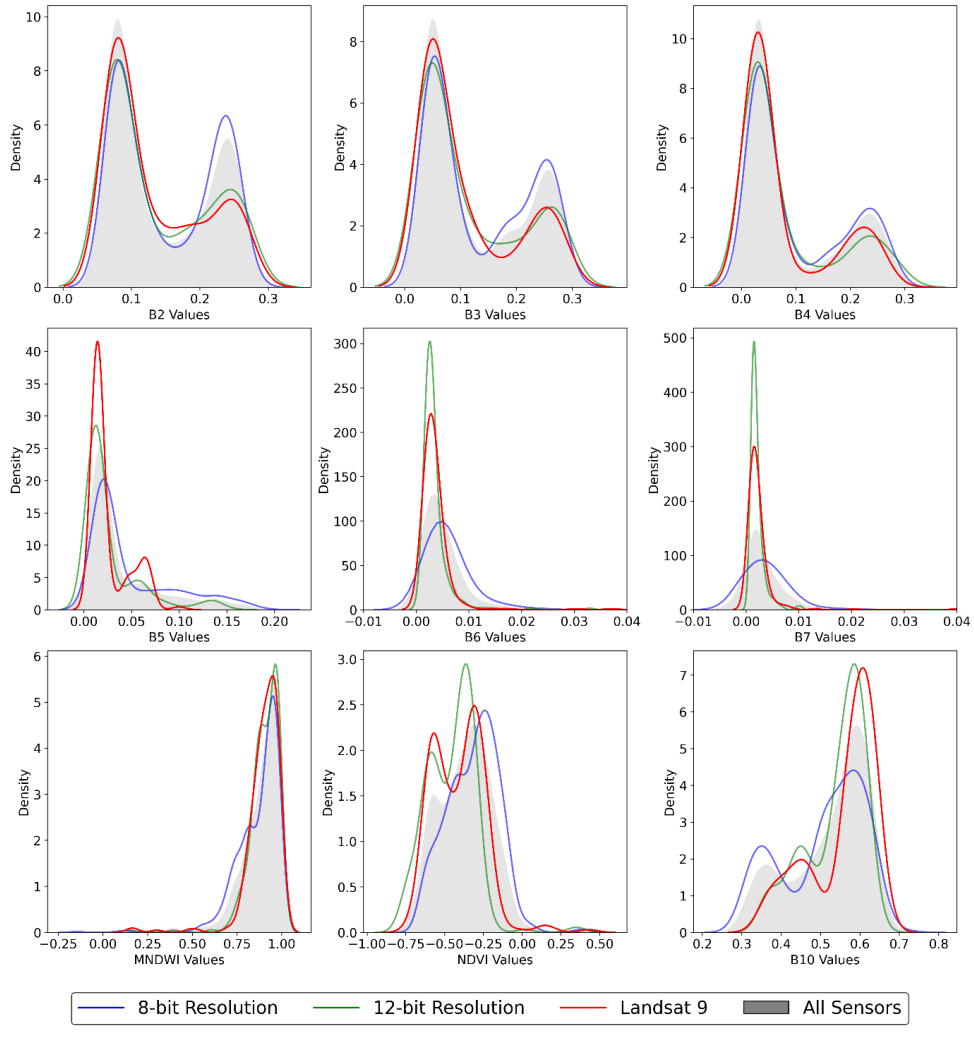


Figure 7 Kernel density estimate (KDE) plot to visualise the distribution of band values for 8 and 12-bit resolution sensors, Landsat 9 and all sensors for water.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

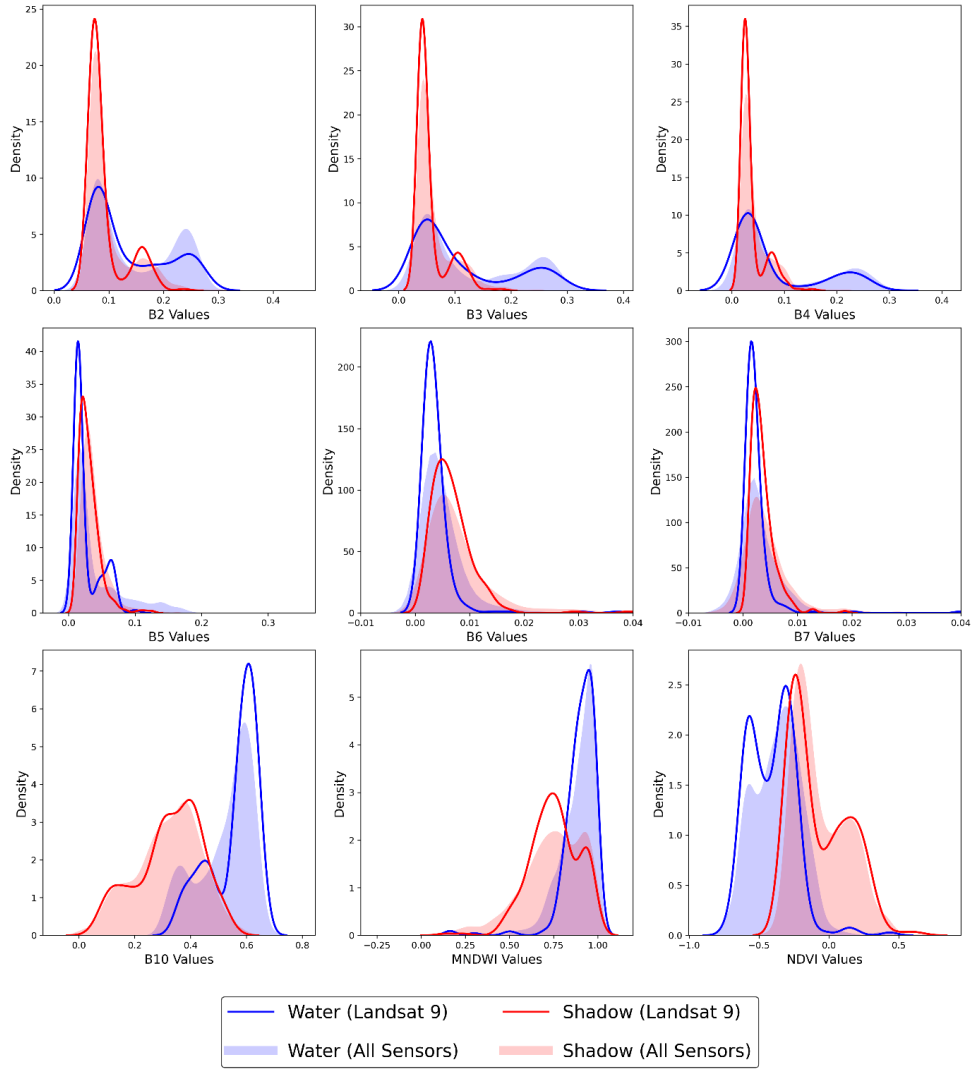


Figure 8 Kernel density estimate (KDE) plot to visualise the distribution of bands and indices values for the different land classes for Landsat 9.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

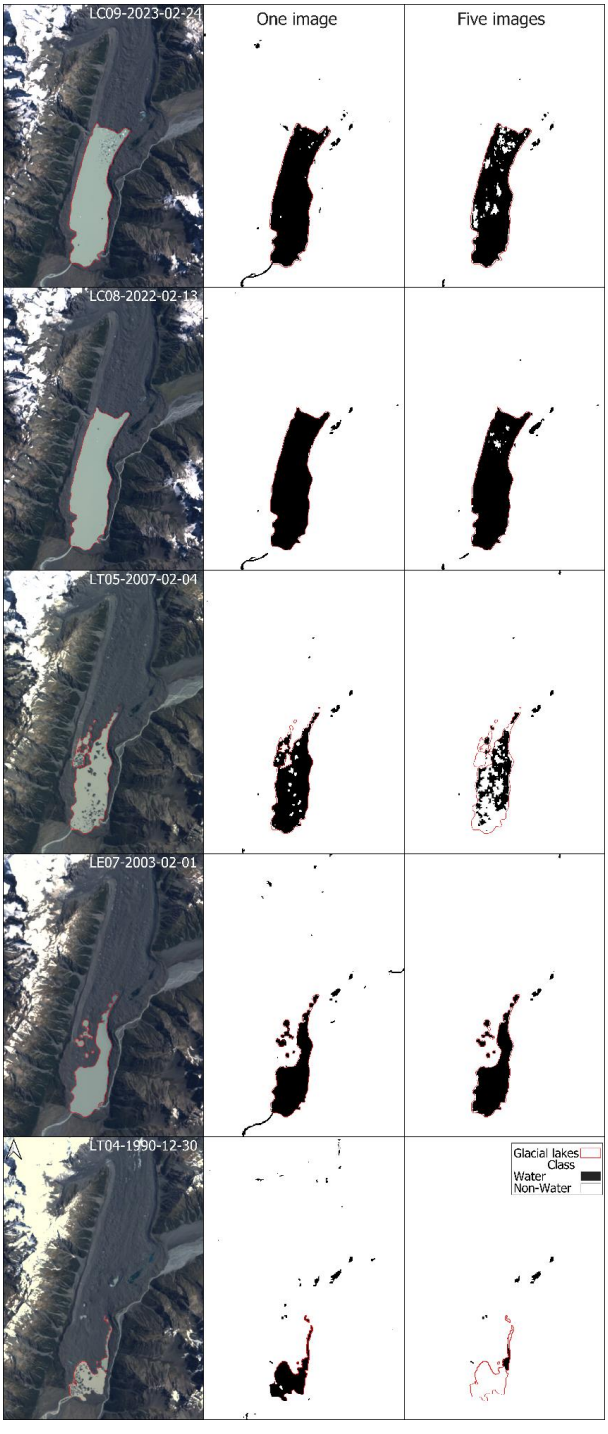


Figure 9 The comparison of five different images with the binary OBIA classification for one image trained against 5 images trained. Landsat 4-9 Collection 2 Tier 1 TOA imagery courtesy of the U.S. Geological Survey.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

4.4 Comparison of water extraction methods and the OBIA classifier

Despite the initial longer processing time required, the OBIA approach compared to other traditional methods of water extraction in GEE reduced the time needed for manually identifying threshold values for different sensors (Ali et al., 2023; Rastner et al., 2014). We applied scene specific thresholding based on the five training images in the collection which gave us large thresholds due to the spectral differences between lakes and between sensors. This additional step of manually inspecting and selecting threshold values was time-consuming. Additionally, the comparison between the OBIA and other water extraction indices (NDWI, MNDWI, AWEI_sh) demonstrates the superiority of the OBIA in delineating water bodies with the addition of bands (Table 7, Figure 10, Figure 11). This is highlighted when using the NDWI index, as shadow and water body values are similar, which can lead to misclassification in the image, a limitation also seen in mapping glacial lakes in the Himalaya (Li and Sheng, 2012). Of the methods tested, MNDWI is the worst performing water index as it was not able to classify dark surfaces correctly due to SWIR band spectral values (Guo et al., 2017). Similarly, the MNDWI and AWEI_sh in this study misclassified snow and shadow, leading to false positives in water detection (Acharya et al., 2018; Feyisa et al., 2014). The superiority of the OBIA is represented by having the highest percentage accuracy of water of 94% whilst the RF PBIA approach, NDWI, MNDWI and AWEI_sh had 74.5%, 68.4%, 56.6% and 56%, respectively.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

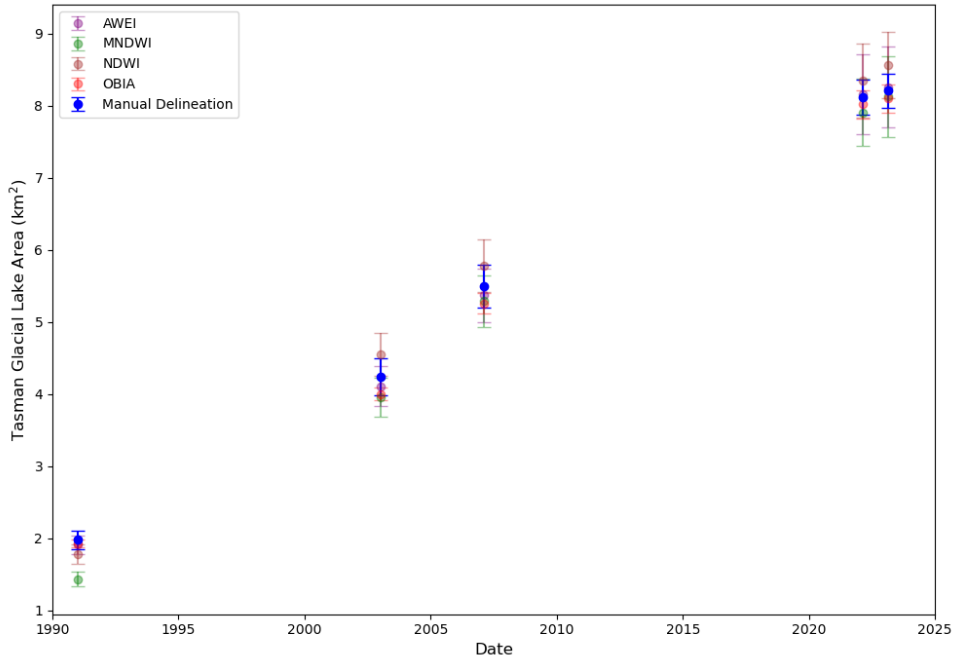


Figure 10 Glacial lake area (km^2) for the Tasman glacial lake using the OBIA, pixel-based indices and manual delineation.

When applying a RF PBIA approach using the same six feature classes and 250 points per class, we find that it produces less accurate results than the OBIA classification. Despite higher accuracies than individual water-based indices, the RF PBIA approach persists in the misclassification of pixels due to the highly diverse spectral properties of differing feature classes in images (Jawak et al., 2022). This is evident between water and shadow where there is a high amount of shadowed and dense vegetation surfaces misclassified as high amounts of speckled pixels, a disadvantage commonly observed in a PBIA approach (Figure 11, Rastner et al., 2014). Manual delineation still remains an accurate way of mapping water bodies however, this method can give inconsistent and generalised results, requiring multiple outlines created to avoid user bias and takes around 100-300 times longer than automated methods (Paul et al., 2013). Our OBIA approach is superior in the delineation of different feature classes in Landsat imagery and can be applied at high accuracy throughout the whole Landsat collection without the need to manually inspect pixel values in individual images.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

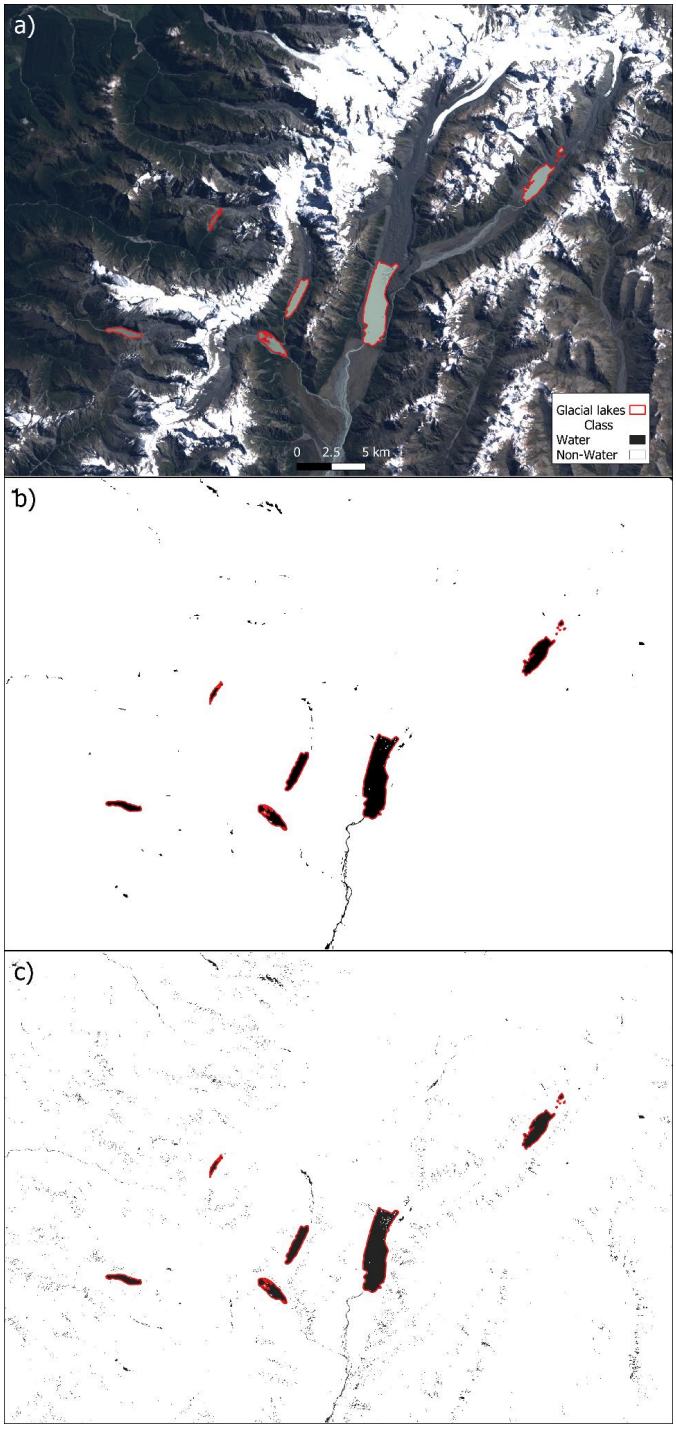


Figure 11 The a) Original image, b) the OBIA classification and c) the binary PBIA classification map. Landsat 9 Collection 2 Tier 1 TOA imagery courtesy of the U.S. Geological Survey.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

5 Conclusions

We ran multiple experiments and comparisons to create and test a new and efficient OBIA approach for mapping glacial lakes across the entire Landsat archive. We find that the best combination for classification and segmentation points was 100-250 (RF, number of points per class) and 4-128 (connectivity, *neighborhoodSize*). We find that using a DEM and topographic features improves the distinction between water bodies and shadow in Landsat images, particularly because of the spectral similarities between the two land classes. Our studies show that the addition of a hillshade and slope provided by a DEM reduced the misclassification 50.7% and eliminated the misclassification of shadow as water in an image.

The inclusion of the MNDWI and NDVI were important for improving the delineation of water and vegetation bodies due to their stable thresholds and can be used as additional indices to aid in the OBIA classifier to delineate land classes in Landsat 4-9 satellite images. This is shown by the improvement in the overall accuracy of the OBIA classification, and the producer's and user's accuracy for water, respectively. Therefore, future studies should consider applying additional data sets and indices to improve the overall accuracy of an OBIA approach.

Our study identifies that training and evaluating one Landsat 9 image produces a rapid and efficient OBIA classifier that can be applied to the whole collection, significantly outperforming the five image OBIA classifier. We found that using a single image from each Landsat 4-7 sensor can produce a rapid and efficient OBIA classifier that can be reproduced successfully for its own image collection. However, due to the complexities of radiometric sensitivity, heterogenous values for bands and indices and temporal changes in land cover throughout the study, applying a five image OBIA classifier to an image collection leads to increased misclassification between land classes and most importantly in water detection.

Our study presents a new and robust tool for glacial lake extraction through OBIA that can be applied to the whole multi-spectral Landsat resulting in an overall accuracy of 94.6%, with a producer's accuracy and user's accuracy of 95.3 % and 95.5% respectively for water. This demonstrates that our OBIA approach can accurately and rapidly delineate glacial lakes on a regional scale, applied to any image in the Landsat era, and is an improvement on the previous work of pixel-based and manual delineation approaches by reducing the time required of manual correction. Our OBIA approach will support the future development of a better understanding of the changes in lake-terminating glaciers and lakes in the Southern Alps of New Zealand - by providing an accurate, rapid and automated method to rapidly extract glacial lakes that will aid in understanding the magnitude of their contribution to past and future regional ice mass loss. Future studies may consider applying an OBIA approach to Sentinel 2 imagery, to improve the number of images studied, due to its high revisit frequency of 5 days that can be combined with Landsat 8 and 9 imageries, with the possibility of monitoring glacial lakes every two to three days.

Code and data availability

The data generated for this study is available at https://github.com/tomosglaciology/GlacialLakeExtractions. The repository contains the figures, scripts and documentation necessary to reproduce the methods and results presented in this manuscript. Our Google Earth Engine script can be accessed directly at: https://code.earthengine.google.com/4ea63fd5a2b413db47dc07c55a4622d4.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Author contributions

TM and RM contributed to methodology development and drafted the original manuscript. RM and PD contributed to reviewing and editing the manuscript.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Disclaimer

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements

Financial support

This research has been funded by the Department of the Economy (DfE) in Northern Ireland.

Review statement

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

References

Achanta, R. and Susstrunk, S.: Superpixels and Polygons Using Simple Non-iterative Clustering, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 4895–4904, https://doi.org/10.1109/CVPR.2017.520, 2017.

Acharya, T. D., Subedi, A., and Lee, D. H.: Evaluation of Water Indices for Surface Water Extraction in a Landsat 8 Scene of Nepal, Sensors, 18, 2580, https://doi.org/10.3390/s18082580, 2018.

Ali, A., Dunlop, P., Coleman, S., Kerr, D., McNabb, R. W., and Noormets, R.: Glacier area changes in Novaya Zemlya from 1986–89 to 2019–21 using object-based image analysis in Google Earth Engine, J. Glaciol., 69, 1305–1316, https://doi.org/10.1017/jog.2023.18, 2023.

Amani, M., Ghorbanian, A., Ahmadi, S. A., Kakooei, M., Moghimi, A., Mirmazloumi, S. M., Moghaddam, S. H. A., Mahdavi, S., Ghahremanloo, M., Parsian, S., Wu, Q., and Brisco, B.: Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 13, 5326–5350, https://doi.org/10.1109/JSTARS.2020.3021052, 2020.

Baumann, S., Anderson, B., Chinn, T., Mackintosh, A., Collier, C., Lorrey, A. M., Rack, W., Purdie, H., and Eaves, S.: Updated inventory of glacier ice in New Zealand based on 2016 satellite imagery, J. Glaciol., 67, 13–26, https://doi.org/10.1017/jog.2020.78, 2021.

Belgiu, M. and Drăguţ, L.: Random forest in remote sensing: A review of applications and future directions, ISPRS J. Photogramm. Remote Sens., 114, 24–31, https://doi.org/10.1016/j.isprsjprs.2016.01.011, 2016.

Bhambri, R., Mehta, M., Dobhal, D. P., and Gupta, A. K.: Glacier lake inventory of Uttarakhand, Wadia Institute of Himalayan Geology, 2015.

Blaschke, T.: Object based image analysis for remote sensing, ISPRS J. Photogramm. Remote Sens., 65, 2–16, https://doi.org/10.1016/J.ISPRSJPRS.2009.06.004, 2010.

Breiman, L., Friedman, J., Olshen, R. A., and Stone, C. J.: Classification and Regression Trees, Chapman and Hall/CRC, New York, 368 pp., https://doi.org/10.1201/9781315139470, 2017.

Bueno, I. T., Acerbi Júnior, F. W., Silveira, E. M. O., Mello, J. M., Carvalho, L. M. T., Gomide, L. R., Withey, K., and Scolforo, J. R. S.: Object-Based Change Detection in the Cerrado Biome Using Landsat Time Series, Remote Sens., 11, 570, https://doi.org/10.3390/rs11050570, 2019.

Capolupo, A., Monterisi, C., and Tarantino, E.: Landsat Images Classification Algorithm (LICA) to Automatically Extract Land Cover Information in Google Earth Engine Environment, Remote Sens., 12, 1201, https://doi.org/10.3390/rs12071201, 2020.

Carrivick, J. L., Sutherland, J. L., Huss, M., Purdie, H., Stringer, C. D., Grimes, M., James, W. H. M., and Lorrey, A. M.: Coincident evolution of glaciers and ice-marginal proglacial lakes across the Southern Alps, New Zealand: Past, present and future, Glob. Planet. Change, 211, 103792, https://doi.org/10.1016/J.GLOPLACHA.2022.103792, 2022.

Castilla, G. and Hay, G. J.: Image objects and geographic objects, in: Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications, edited by: Blaschke, T., Lang, S., and Hay, G. J., Springer, Berlin, Heidelberg, 91–110, https://doi.org/10.1007/978-3-540-77058-9_5, 2008.

Chen, F., Zhang, M., Tian, B., and Li, Z.: Extraction of Glacial Lake Outlines in Tibet Plateau Using Landsat 8 Imagery and Google Earth Engine, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 10, 4002–4009, https://doi.org/10.1109/JSTARS.2017.2705718, 2017.

Chen, Z. and Zhao, S.: Automatic monitoring of surface water dynamics using Sentinel-1 and Sentinel-2 data with Google Earth Engine, Int. J. Appl. Earth Obs. Geoinformation, 113, 103010, https://doi.org/10.1016/j.jag.2022.103010, 2022.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Chinn, T., Fitzharris, B. B., Willsman, A., and Salinger, M. J.: Annual ice volume changes 1976–2008 for the New Zealand Southern Alps, Glob. Planet. Change, 92–93, 105–118, https://doi.org/10.1016/j.gloplacha.2012.04.002, 2012.

Chinn, T. J.: Distribution of the glacial water resources of New Zealand, J. Hydrol. N. Z., 40, 139–187, http://www.jstor.org/stable/43922047, 2001.

Chinn, T. J., Heydenrych, C., and Salinger, M. J.: Use of the ELA as a practical method of monitoring glacier response to climate in New Zealand's Southern Alps, J. Glaciol., 51, 85–95, https://doi.org/10.3189/172756505781829593, 2005.

Congalton, R. G.: A review of assessing the accuracy of classifications of remotely sensed data, Remote Sens. Environ., 37, 35–46, https://doi.org/10.1016/0034-4257(91)90048-B, 1991.

Congedo, L.: Semi-Automatic Classification Plugin for QGIS, Sapienza University, Rome, 2013.

Congedo, L.: Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS, J. Open Source Softw., 6, 3172, https://doi.org/10.21105/joss.03172, 2021.

Crawford, C. J., Roy, D. P., Arab, S., Barnes, C., Vermote, E., Hulley, G., Gerace, A., Choate, M., Engebretson, C., Micijevic, E., Schmidt, G., Anderson, C., Anderson, M., Bouchard, M., Cook, B., Dittmeier, R., Howard, D., Jenkerson, C., Kim, M., Kleyians, T., Maiersperger, T., Mueller, C., Neigh, C., Owen, L., Page, B., Pahlevan, N., Rengarajan, R., Roger, J.-C., Sayler, K., Scaramuzza, P., Skakun, S., Yan, L., Zhang, H. K., Zhu, Z., and Zahn, S.: The 50-year Landsat collection 2 archive, Sci. Remote Sens., 8, 100103, https://doi.org/10.1016/j.srs.2023.100103, 2023.

Csillik, O.: Fast segmentation and classification of very high resolution remote sensing data using SLIC superpixels, Remote Sens., 9, 243, 2017.

Dou, X., Fan, X., Wang, X., Yunus, A. P., Xiong, J., Tang, R., Lovati, M., van Westen, C., and Xu, Q.: Spatio-Temporal Evolution of Glacial Lakes in the Tibetan Plateau over the Past 30 Years, Remote Sens., 15, 416, https://doi.org/10.3390/rs15020416, 2023.

Dykes, R. C.: A multi-parameter study of iceberg calving and the retreat of Haupapa/Tasman Glacier, South Island, New Zealand : a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Geography at Massey University, Palmerston North, New Zealand, Doctoral, Massey University, 2013.

Dykes, R. C., Brook, M. S., Robertson, C. M., and Fuller, I. C.: Twenty-First Century Calving Retreat of Tasman Glacier, Southern Alps, New Zealand, Arct. Antarct. Alp. Res., 43, 1–10, https://doi.org/10.1657/1938-4246-43.1.1, 2011.

Earth Engine Data Catalog | Google for Developers: https://developers.google.com/earth-engine/datasets, last access: 12 February 2025.

Fahnestock, M., Scambos, T., Moon, T., Gardner, A., Haran, T., and Klinger, M.: Rapid large-area mapping of ice flow using Landsat 8, Remote Sens. Environ., 185, 84–94, https://doi.org/10.1016/J.RSE.2015.11.023, 2016.

Feyisa, G. L., Meilby, H., Fensholt, R., and Proud, S. R.: Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery, Remote Sens. Environ., 140, 23–35, https://doi.org/10.1016/J.RSE.2013.08.029, 2014.

Foody, G. M. and Mathur, A.: Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification, Remote Sens. Environ., 93, 107–117, https://doi.org/10.1016/j.rse.2004.06.017, 2004.

Garcia, R.: Mapping the Retreat of the Debris-Covered Tasman Glacier in the Aoraki-Mount Cook National Park, New Zealand, 2022.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote Sens. Environ., 202, 18–27, https://doi.org/10.1016/j.rse.2017.06.031, 2017.

Gujrati, A. and Jha, V. B.: SURFACE WATER DYNAMICS OF INLAND WATER BODIES OF INDIA USING GOOGLE EARTH ENGINE, ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci., IV–5, 467–472, https://doi.org/10.5194/isprs-annals-IV-5-467-2018, 2018.

Guo, Q., Pu, R., Li, J., and Cheng, J.: A weighted normalized difference water index for water extraction using Landsat imagery, Int. J. Remote Sens., 38, 5430–5445, https://doi.org/10.1080/01431161.2017.1341667, 2017.

Hall, D. K., Riggs, G. A., and Salomonson, V. V.: Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data, Remote Sens. Environ., 54, 127–140, https://doi.org/10.1016/0034-4257(95)00137-P, 1995.

Hall, D. K., Bayr, K. J., Schöner, W., Bindschadler, R. A., and Chien, J. Y. L.: Consideration of the errors inherent in mapping historical glacier positions in Austria from the ground and space (1893–2001), Remote Sens. Environ., 86, 566–577, https://doi.org/10.1016/S0034-4257(03)00134-2, 2003.

Henderson, R. D. and Thompson, S. M.: Extreme Rainfalls in the Southern Alps of New Zealand, J. Hydrol. N. Z., 38, 309–330, 1999.

Hochstein, M. P., Claridge, D., Henrys, S. A., Pyne, A., Nobes, D. C., and Leary, S. F.: Downwasting of the Tasman Glacier, South Island, New Zealand: Changes in the terminus region between 1971 and 1993, N. Z. J. Geol. Geophys., 38, 1–16, https://doi.org/10.1080/00288306.1995.9514635, 1995.

Hossain, M. D. and Chen, D.: Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective, ISPRS J. Photogramm. Remote Sens., 150, 115–134, https://doi.org/10.1016/j.isprsjprs.2019.02.009, 2019.

Irons, J. R., Dwyer, J. L., and Barsi, J. A.: The next Landsat satellite: The Landsat Data Continuity Mission, Remote Sens. Environ., 122, 11–21, https://doi.org/10.1016/j.rse.2011.08.026, 2012.

Jawak, S. D. and Luis, A. J.: A Semiautomatic Extraction of Antarctic Lake Features Using Worldview-2 Imagery, Photogramm. Eng. Remote Sens., 80, 939–952, https://doi.org/10.14358/PERS.80.10.939, 2014.

Jawak, S. D., Wankhede, S. F., Luis, A. J., and Balakrishna, K.: 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 Sens., 14, 6311, https://doi.org/10.3390/rs14246311, 2022.

Jeong, S. and Howat, I. M.: Performance of Landsat 8 Operational Land Imager for mapping ice sheet velocity, Remote Sens. Environ., 170, 90–101, https://doi.org/10.1016/j.rse.2015.08.023, 2015.

Ji, L., Zhang, L., and Wylie, B.: Analysis of Dynamic Thresholds for the Normalized Difference Water Index, Photogramm. Eng. Remote Sens., 75, 1307–1317, https://doi.org/10.14358/PERS.75.11.1307, 2009.

Jiang, Z., Huete, A. R., Didan, K., and Miura, T.: Development of a two-band enhanced vegetation index without a blue band, Remote Sens. Environ., 112, 3833–3845, https://doi.org/10.1016/j.rse.2008.06.006, 2008.

Karakuş, P.: Object Based Classification in Google Earth Engine Combining SNIC and Machine Learning Methods (Case Study: Lake Köyceğiz), Turk. J. Remote Sens. GIS, 125–137, https://doi.org/10.48123/rsgis.1411380, 2024.

Kingslake, J., Ely, J. C., Das, I., and Bell, R. E.: Widespread movement of meltwater onto and across Antarctic ice shelves, Nature, 544, 349–352, https://doi.org/10.1038/nature22049, 2017.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Li, J. and Sheng, Y.: An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models: a case study in the Himalayas, Int. J. Remote Sens., 33, 5194–5213, https://doi.org/10.1080/01431161.2012.657370, 2012.

Lillesand, T., Kiefer, R. W., and Chipman, J.: Remote Sensing and Image Interpretation, John Wiley & Sons, 775 pp., 2015.

LINZ Data Service: LINZ NZ 8m Digital Elevation Model, 2012.

Lu, Y., Zhang, Z., Shangguan, D., and Yang, J.: Novel Machine Learning Method Integrating Ensemble Learning and Deep Learning for Mapping Debris-Covered Glaciers, Remote Sens., 13, 2595, https://doi.org/10.3390/rs13132595, 2021.

Mackintosh, A. N., Anderson, B. M., Lorrey, A. M., Renwick, J. A., Frei, P., and Dean, S. M.: Regional cooling caused recent New Zealand glacier advances in a period of global warming, Nat. Commun., 8, 14202, https://doi.org/10.1038/ncomms14202, 2017.

Markham, B. L., Storey, J. C., Williams, D. L., and Irons, J. R.: Landsat sensor performance: history and current status, IEEE Trans. Geosci. Remote Sens., 42, 2691–2694, https://doi.org/10.1109/TGRS.2004.840720, 2004.

Masek, J. G., Wulder, M. A., Markham, B., McCorkel, J., Crawford, C. J., Storey, J., and Jenstrom, D. T.: Landsat 9: Empowering open science and applications through continuity, Remote Sens. Environ., 248, 111968, https://doi.org/10.1016/j.rse.2020.111968, 2020.

McFeeters, S. K.: The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, Int. J. Remote Sens., 17, 1425–1432, https://doi.org/10.1080/01431169608948714, 1996.

Millard, K. and Richardson, M.: On the Importance of Training Data Sample Selection in Random Forest Image Classification: A Case Study in Peatland Ecosystem Mapping, Remote Sens., 7, 8489–8515, https://doi.org/10.3390/rs70708489, 2015.

Mitkari, K. V., Arora, M. K., and Tiwari, R. K.: Extraction of Glacial Lakes in Gangotri Glacier Using Object-Based Image Analysis, IEEE J. Sel. Top. Appl. EARTH Obs. REMOTE Sens., 10, https://doi.org/10.1109/JSTARS.2017.2727506, 2017.

Mitkari, K. V., Arora, M. K., Tiwari, R. K., Sofat, S., Gusain, H. S., and Tiwari, S. P.: Large-Scale Debris Cover Glacier Mapping Using Multisource Object-Based Image Analysis Approach, Remote Sens., 14, 3202, https://doi.org/10.3390/rs14133202, 2022.

Mondejar, J. P. and Tongco, A. F.: Near infrared band of Landsat 8 as water index: a case study around Cordova and Lapu-Lapu City, Cebu, Philippines, Sustain. Environ. Res., 29, 16, https://doi.org/10.1186/s42834-019-0016-5, 2019.

Mouginot, J., Rignot, E., Scheuchl, B., and Millan, R.: Comprehensive Annual Ice Sheet Velocity Mapping Using Landsat-8, Sentinel-1, and RADARSAT-2 Data, Remote Sens., 9, 364, https://doi.org/10.3390/rs9040364, 2017.

Myneni, R. B., Hall, F. G., Sellers, P. J., and Marshak, A. L.: The Interpretation of Spectral Vegetation Indexes, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, https://doi.org/10.1109/TGRS.1995.8746029, 1995.

Nie, Y., Liu, Q., Wang, J., Zhang, Y., Sheng, Y., and Liu, S.: An inventory of historical glacial lake outburst floods in the Himalayas based on remote sensing observations and geomorphological analysis, Geomorphology, 308, 91–106, https://doi.org/10.1016/J.GEOMORPH.2018.02.002, 2018.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., and Wulder, M. A.: Good practices for estimating area and assessing accuracy of land change, Remote Sens. Environ., 148, 42–57, https://doi.org/10.1016/j.rse.2014.02.015, 2014.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Oshiro, T. M., Perez, P. S., and Baranauskas, J. A.: How Many Trees in a Random Forest?, in: Machine Learning and Data Mining in Pattern Recognition, Berlin, Heidelberg, 154–168, https://doi.org/10.1007/978-3-642-31537-4_13, 2012.

Pal, M.: Random forest classifier for remote sensing classification, Int. J. Remote Sens., 26, 217–222, https://doi.org/10.1080/01431160412331269698, 2005.

Pal, M. and Mather, P. M.: An assessment of the effectiveness of decision tree methods for land cover classification, Remote Sens. Environ., 86, 554–565, https://doi.org/10.1016/S0034-4257(03)00132-9, 2003.

Paul, F., Barrand, N. E., Baumann, S., Berthier, E., Bolch, T., Casey, K., Frey, H., Joshi, S. P., Konovalov, V., Bris, R. L., Mölg, N., Nosenko, G., Nuth, C., Pope, A., Racoviteanu, A., Rastner, P., Raup, B., Scharrer, K., Steffen, S., and Winsvold, S.: On the accuracy of glacier outlines derived from remote-sensing data, Ann. Glaciol., 54, 171–182, https://doi.org/10.3189/2013AoG63A296, 2013.

Paul, F., Baumann, S., Anderson, B., and Rastner, P.: Deriving a year 2000 glacier inventory for New Zealand from the existing 2016 inventory, Ann. Glaciol., 1–11, https://doi.org/10.1017/aog.2023.20, 2023.

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., Tucker, C. J., and Stenseth, N. C.: Using the satellite-derived NDVI to assess ecological responses to environmental change, Trends Ecol. Evol., 20, 503–510, https://doi.org/10.1016/j.tree.2005.05.011, 2005.

Prathusha, P. and Jyothi, S.: A Novel Edge Detection Algorithm for Fast and Efficient Image Segmentation, in: Data Engineering and Intelligent Computing, Singapore, 283–291, https://doi.org/10.1007/978-981-10-3223-3_26, 2018.

Probst, P. and Boulesteix, A.-L.: To tune or not to tune the number of trees in random forest?, J. Mach. Learn. Res. 18 2018 1-18, https://doi.org/10.48550/arXiv.1705.05654, 2018.

Qayyum, N., Ghuffar, S., Ahmad, H. M., Yousaf, A., and Shahid, I.: Glacial Lakes Mapping Using Multi Satellite PlanetScope Imagery and Deep Learning, ISPRS Int. J. Geo-Inf., 9, 560, https://doi.org/10.3390/ijgi9100560, 2020.

QGIS geographic information system: http://www.qgis.org/.

Rama Rao, N., Garg, P. K., and Ghosh, S. K.: Evaluation of radiometric resolution on land use/land cover mapping in an agricultural area, Int. J. Remote Sens., 28, 443–450, https://doi.org/10.1080/01431160600733181, 2007.

Rastner, P., Bolch, T., Notarnicola, C., and Paul, F.: A Comparison of Pixel- and Object-Based Glacier Classification With Optical Satellite Images, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 7, 853, https://doi.org/10.1109/JSTARS.2013.2274668, 2014.

Robertson, C. M., Benn, D. I., Brook, M. S., Fuller, Ian C., and Holt, K. A.: Subaqueous calving margin morphology at Mueller, Hooker and Tasman glaciers in Aoraki/Mount Cook National Park, New Zealand, J. Glaciol., 58, 1037–1046, https://doi.org/10.3189/2012JoG12J048, 2012.

Röhl, K.: Thermo-erosional notch development at fresh-water-calving Tasman Glacier, New Zealand, J. Glaciol., 52, 203–213, https://doi.org/10.3189/172756506781828773, 2006.

Röhl, K.: Characteristics and evolution of supraglacial ponds on debris-covered Tasman Glacier, New Zealand, J. Glaciol., 54, 867–880, https://doi.org/10.3189/002214308787779861, 2008.

Sahu, R. and Gupta, R. D.: CONCEPTUAL FRAMEWORK OF COMBINED PIXEL AND OBJECT-BASED METHOD FOR DELINEATION OF DEBRIS-COVERED GLACIERS, ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci., IV–5, 173–180, https://doi.org/10.5194/isprs-annals-IV-5-173-2018, 2018.

Shafizadeh-Moghadam, H., Khazaei, M., Alavipanah, S. K., and Weng, Q.: Google Earth Engine for large-scale land use and land cover mapping: an object-based classification approach using spectral, textural and

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

topographical factors, GIScience Remote Sens., 58, 914–928, https://doi.org/10.1080/15481603.2021.1947623, 2021.

Shugar, Burr, A., Haritashya, U. K., Kargel, J. S., Watson, C. S., Kennedy, M. C., Bevington, A. R., Betts, R. A., Harrison, S., and Strattman, K.: Rapid worldwide growth of glacial lakes since 1990, Nat. Clim. Change, 10, 939–945, https://doi.org/10.1038/s41558-020-0855-4, 2020.

Singh, K. V., Setia, R., Sahoo, S., Prasad, A., and Pateriya, B.: Evaluation of NDWI and MNDWI for assessment of waterlogging by integrating digital elevation model and groundwater level, Geocarto Int., 30, 650–661, https://doi.org/10.1080/10106049.2014.965757, 2015.

Spasojević, S. and Clayton, R. W.: Crustal structure and apparent tectonic underplating from receiver function analysis in South Island, New Zealand, J. Geophys. Res. Solid Earth, 113, https://doi.org/10.1029/2007JB005166, 2008.

Sturman, A. P. and Tapper, N. J.: The weather and climate of Australia and New Zealand, Oxford University Press, USA, 1996.

Su, T. and Zhang, S.: Local and global evaluation for remote sensing image segmentation, ISPRS J. Photogramm. Remote Sens., 130, 256–276, https://doi.org/10.1016/j.isprsjprs.2017.06.003, 2017.

Tang, Q., Zhang, G., Yao, T., Wieland, M., Liu, L., and Kaushik, S.: Automatic extraction of glacial lakes from Landsat imagery using deep learning across the Third Pole region, Remote Sens. Environ., 315, 114413, https://doi.org/10.1016/j.rse.2024.114413, 2024.

Tassi, A. and Vizzari, M.: Object-Oriented LULC Classification in Google Earth Engine Combining SNIC, GLCM, and Machine Learning Algorithms, Remote Sens., 12, 3776, https://doi.org/10.3390/rs12223776, 2020.

Tian, B., Li, Z., Zhang, M., Huang, L., Qiu, Y., Li, Z., and Tang, P.: Mapping Thermokarst Lakes on the Qinghai–Tibet Plateau Using Nonlocal Active Contours in Chinese GaoFen-2 Multispectral Imagery, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 10, 1687–1700, https://doi.org/10.1109/JSTARS.2017.2666787, 2017.

Tu, Y., Chen, B., Zhang, T., and Xu, B.: Regional mapping of essential urban land use categories in China: A segmentation-based approach, Remote Sens., 12, 1058, https://doi.org/10.3390/rs12071058, 2020.

Tucker, C. J.: Radiometric resolution for monitoring vegetation How many bits are needed?, Int. J. Remote Sens., https://doi.org/10.1080/01431168008948235, 1980.

Veh, G., Korup, O., Roessner, S., and Walz, A.: Detecting Himalayan glacial lake outburst floods from Landsat time series, Remote Sens. Environ., 207, 84–97, https://doi.org/10.1016/j.rse.2017.12.025, 2018.

Verde, N., Mallinis, G., Tsakiri-Strati, M., Georgiadis, C., and Patias, P.: Assessment of Radiometric Resolution Impact on Remote Sensing Data Classification Accuracy, Remote Sens., 10, 1267, https://doi.org/10.3390/rs10081267, 2018.

Wang, W., Xiang, Y., Gao, Y., Lu, A., and Yao, T.: Rapid expansion of glacial lakes caused by climate and glacier retreat in the Central Himalayas, Hydrol. Process., 29, 859–874, https://doi.org/10.1002/hyp.10199, 2015.

Wang, X., Ding, Y., Liu, S., Jiang, L., Wu, K., Jiang, Z., and Guo, W.: Changes of glacial lakes and implications in Tian Shan, central Asia, based on remote sensing data from 1990 to 2010, Environ. Res. Lett., 8, 044052, https://doi.org/10.1088/1748-9326/8/4/044052, 2013.

Wang, X., Guo, X., Yang, C., Liu, Q., Wei, J., Zhang, Y., Liu, S., Zhang, Y., Jiang, Z., and Tang, Z.: Glacial lake inventory of high-mountain Asia in 1990 and 2018 derived from Landsat images, Earth Syst. Sci. Data, 12, 2169–2182, https://doi.org/10.5194/essd-12-2169-2020, 2020.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

878 Wangchuk, S. and Bolch, T.: Mapping of glacial lakes using Sentinel-1 and Sentinel-2 data and a random forest
879 classifier: Strengths and challenges, Sci. Remote Sens., 2, 100008, https://doi.org/10.1016/j.srs.2020.100008,
880 2020.

881 Warren, C. R. and Kirkbride, M. P.: Temperature and bathymetry of ice-contact lakes in Mount Cook National
882 Park, New Zealand, N. Z. J. Geol. Geophys., 41, 133–143, https://doi.org/10.1080/00288306.1998.9514797,
883 1998.

884 Weekley, D. and Li, X.: Tracking lake surface elevations with proportional hypsometric relationships, Landsat
885 imagery, and multiple DEMs, Water Resour. Res., 57, e2020WR027666,
886 https://doi.org/10.1029/2020WR027666, 2021.

887 Wessels, R. L., Kargel, J. S., and Kieffer, H. H.: ASTER measurement of supraglacial lakes in the Mount
888 Everest region of the Himalaya, Ann. Glaciol., 34, 399–408, https://doi.org/10.3189/172756402781817545,
889 2002.

890 Whiteside, T. G., Boggs, G. S., and Maier, S. W.: Comparing object-based and pixel-based classifications for
891 mapping savannas, Int. J. Appl. Earth Obs. Geoinformation, 13, 884–893,
892 https://doi.org/10.1016/J.JAG.2011.06.008, 2011.

893 Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., Allen, R. G.,
894 Anderson, M. C., Belward, A. S., Cohen, W. B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla,
895 T., Hipple, J. D., Hostert, P., Hughes, M. J., Huntington, J., Johnson, D. M., Kennedy, R., Kilic, A., Li, Z.,
896 Lymburner, L., McCorkel, J., Pahlevan, N., Scambos, T. A., Schaaf, C., Schott, J. R., Sheng, Y., Storey, J.,
897 Vermote, E., Vogelmann, J., White, J. C., Wynne, R. H., and Zhu, Z.: Current status of Landsat program,
898 science, and applications, Remote Sens. Environ., 225, 127–147, https://doi.org/10.1016/j.rse.2019.02.015,
899 2019.

900 Wulder, M. A., Roy, D. P., Radeloff, V. C., Loveland, T. R., Anderson, M. C., Johnson, D. M., Healey, S., Zhu,
901 Z., Scambos, T. A., Pahlevan, N., Hansen, M., Gorelick, N., Crawford, C. J., Masek, J. G., Hermosilla, T.,
902 White, J. C., Belward, A. S., Schaaf, C., Woodcock, C. E., Huntington, J. L., Lymburner, L., Hostert, P., Gao, F.,
903 Lyapustin, A., Pekel, J.-F., Strobl, P., and Cook, B. D.: Fifty years of Landsat science and impacts, Remote Sens.
904 Environ., 280, 113195, https://doi.org/10.1016/j.rse.2022.113195, 2022.

905 Xia, H., Zhao, J., Qin, Y., Yang, J., Cui, Y., Song, H., Ma, L., Jin, N., and Meng, Q.: Changes in Water Surface
906 Area during 1989–2017 in the Huai River Basin using Landsat Data and Google Earth Engine, Remote Sens.,
907 11, 1824, https://doi.org/10.3390/rs11151824, 2019.

908 Xie, Z., ShangGuan, D., Zhang, S., Ding, Y., and Liu, S.: Index for hazard of Glacier Lake Outburst flood of
909 Lake Merzbacher by satellite-based monitoring of lake area and ice cover, Glob. Planet. Change, 107, 229–237,
910 https://doi.org/10.1016/j.gloplacha.2012.05.025, 2013.

911 Xu, H.: Modification of normalised difference water index (NDWI) to enhance open water features in remotely
912 sensed imagery, Int. J. Remote Sens., 27, 3025–3033, https://doi.org/10.1080/01431160600589179, 2006.

913 Xu, M., Watanachaturaporn, P., Varshney, P. K., and Arora, M. K.: Decision tree regression for soft classification
914 of remote sensing data, Remote Sens. Environ., 97, 322–336, https://doi.org/10.1016/j.rse.2005.05.008, 2005.

915 Yan, D., Huang, C., Ma, N., and Zhang, Y.: Improved Landsat-Based Water and Snow Indices for Extracting
916 Lake and Snow Cover/Glacier in the Tibetan Plateau, https://doi.org/10.3390/w12051339, 2020.

917 Yang, C., Wu, G., Ding, K., Shi, T., Li, Q., and Wang, J.: Improving Land Use/Land Cover Classification by
918 Integrating Pixel Unmixing and Decision Tree Methods, Remote Sens., 9, 1222,
919 https://doi.org/10.3390/rs9121222, 2017.

920 Yang, C.-C., Prasher, S. O., Enright, P., Madramootoo, C., Burgess, M., Goel, P. K., and Callum, I.: Application
921 of decision tree technology for image classification using remote sensing data, Agric. Syst., 76, 1101–1117,
922 https://doi.org/10.1016/S0308-521X(02)00051-3, 2003.

https://doi.org/10.5194/egusphere-2026-2033
Preprint. Discussion started: 8 May 2026

Yang, L., Wang, L., Abubakar, G. A., and Huang, J.: High-resolution rice mapping based on SNIC segmentation and multi-source remote sensing images, Remote Sens., 13, 1148, https://doi.org/10.3390/rs13061148, 2021.

Yue, L., Li, B., Zhu, S., Yuan, Q., and Shen, H.: A fully automatic and high-accuracy surface water mapping framework on Google Earth Engine using Landsat time-series, Int. J. Digit. Earth, 16, 210–233, https://doi.org/10.1080/17538947.2023.2166606, 2023.

Zhang, G., Yao, T., Xie, H., Wang, W., and Yang, W.: An inventory of glacial lakes in the Third Pole region and their changes in response to global warming, Glob. Planet. Change, 131, 148–157, https://doi.org/10.1016/J.GLOPLACHA.2015.05.013, 2015.

Zhang, M., Chen, F., Guo, H., Yi, L., Zeng, J., and Li, B.: Glacial Lake Area Changes in High Mountain Asia during 1990–2020 Using Satellite Remote Sensing, Research, 2022, https://doi.org/10.34133/2022/9821275, 2022.

Zhao, H., Chen, F., and Zhang, M.: A Systematic Extraction Approach for Mapping Glacial Lakes in High Mountain Regions of Asia, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 11, 2788–2799, https://doi.org/10.1109/JSTARS.2018.2846551, 2018.