MapBiomas Water - Ecuador

GLACIERS General Manual

Algorithm Theoretical Basis Document (ATBD)

Collection 3, Version 1

Ecuador team:

Wagner Holguín, Juan Espinosa, Fabricio Garcés.

October 2025

Table of contents

1. Introduction

Glacier mapping is difficult and hazardous due to the remoteness and inaccessibility of the terrain, as well as the challenges involved in conducting extensive fieldwork (Racoviteanu et al., 2009). In this context, data from satellite-mounted sensors, acquired through remote sensing tools, provide valuable information on glaciers and associated landforms. It is important to note that the precise selection of spectral bands is crucial for mapping glacial features (Huang et al., 2021). Glacial landscapes in the Amazon and Ecuador basin, within the Andes biome, have been delineated using selected bands from Landsat program data (Philip & Ravindran, 1998). The vast majority of high-mountain glaciers have experienced accelerated reduction in recent decades (Baťka et al., 2020; Turpo Cayo et al., 2022).

Mountain glaciers, particularly in the tropics, are currently undergoing a rapid decline in surface area (Turpo Cayo et al., 2022) and are considered reliable indicators of climate change. This is due to their relatively fast response time to disturbances in climatic variables such as precipitation, air temperature, and atmospheric humidity (Kaser & Osmaston, 2002). Many tropical glaciers, such as those in Ecuador, Peru, and Bolivia, act as regulators of water resources against reduced precipitation during the dry season. Over 99% of tropical glaciers are located in the Andes of South America, including Venezuela, Colombia, Ecuador, Peru, Bolivia, Chile, and Argentina. Much of the previous research has been conducted using remote sensing data, as the difficult terrain presents challenges for carrying out extensive fieldwork (Veettil & Kamp, 2019).

This document outlines the methodology applied for mapping tropical glaciers in the continental territory of Ecuador for MapBiomas Collection 3. The work was based on the methodology from MapBiomas Amazonia Collection 6 and other references (Turpo Cayo et al., 2022), and consisted of adding the year 2024 to the historical data series. The complete process was divided into 6 stages (Figure 1).

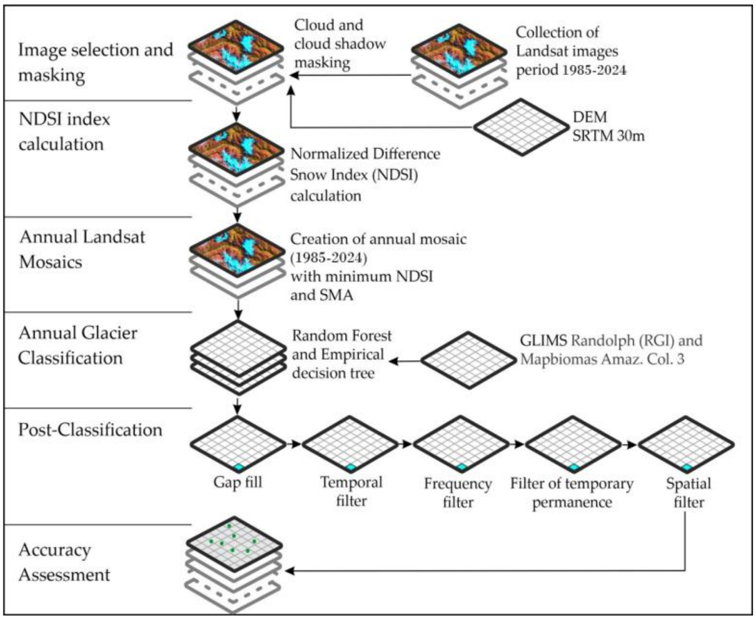

Figure 1: Methodological flowchart for obtaining glacier coverage maps

1.1. Study area

The study area (Figure 2) was defined based on the Randolph Glacier Inventory (RGI Consortium, 2017, p. 0), which encompasses the entire Andean region of Ecuador. Based on this, glacierized zones were delineated as areas extending 500 meters below the altitude of the glaciers detected in 1985, which was considered the baseline for classification.

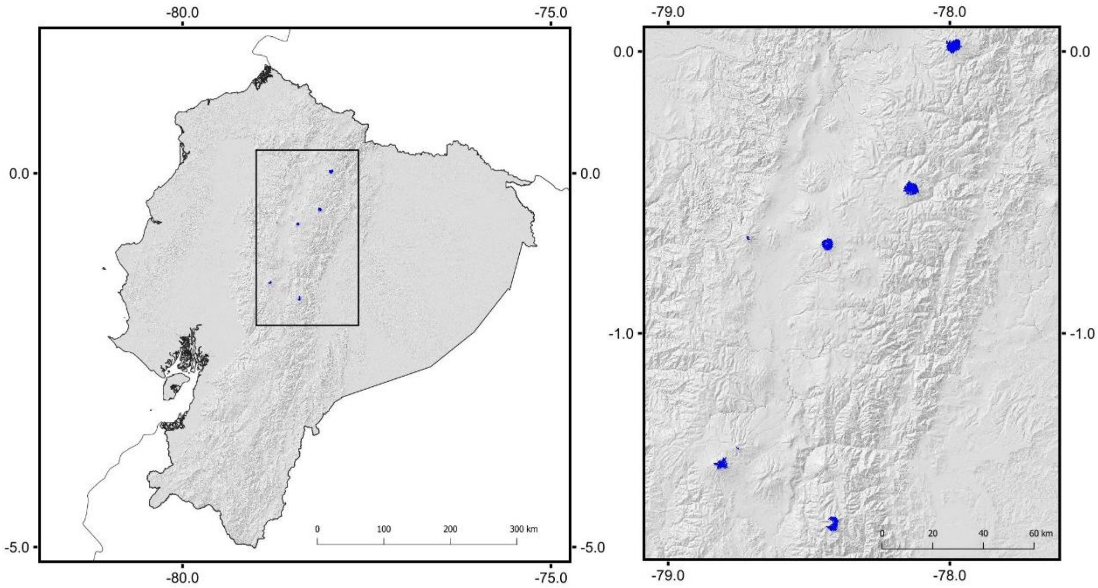

Figure 2: Study area

2. Image mosaics

The classification for the cross-cutting theme "Glaciers" utilized Landsat image mosaics generated specifically for glacier mapping. These mosaics incorporated images with the minimum annual glacier area, based on the pixel with the minimum NDSI quality (Figure 3).

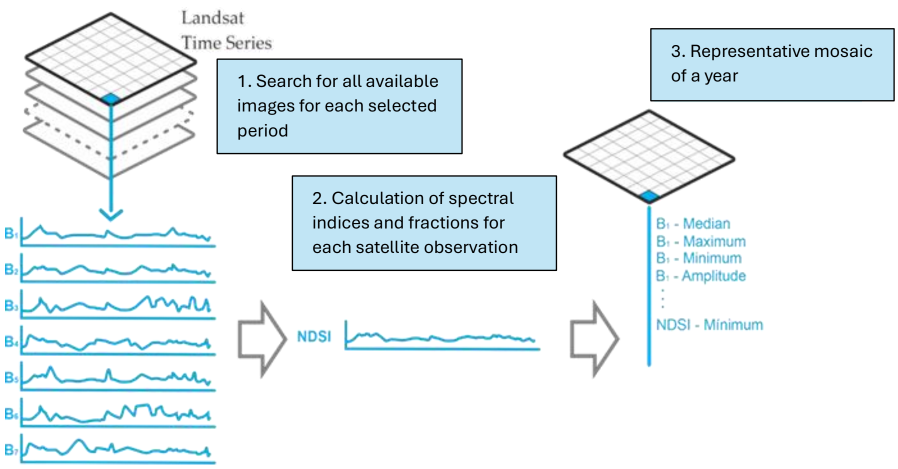

Figure 3: Creation of annual mosaics for Glaciers

3. Classification

The classification of the Landsat mosaics was conducted entirely on the Google Earth Engine platform, based on an empirical tree (Figure 4).

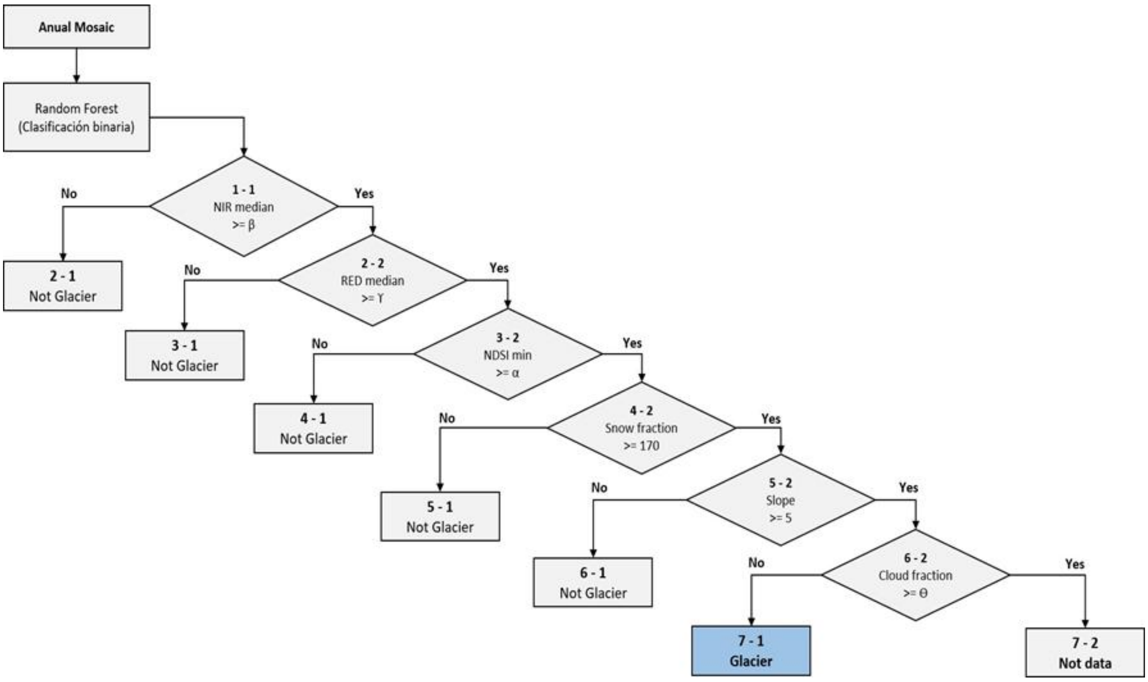

Figure 4: Empirical tree combined with Random Forest for glacier classification.

The glacier classification system implements a hybrid approach that combines machine learning with specific decision rules for different satellite sensors. The classification process begins with the generation of annual, cloud-free mosaics for each study year, using Landsat images from different sensors (L5, L7, L8, and L9) according to the time period.

The Random Forest classification is refined using a set of hierarchical decision rules specific to each sensor type.

- For Landsat 5 and Landsat 7, Near-Infrared (NIR) median reflectance thresholds (1-1) are applied where the values are close to 0.2114.
- For Landsat 8 and Landsat 9, these values are close to 0.1730.
- For the RED median reflectance thresholds (2-2), the estimated value for Landsat 5 and Landsat 7 is 0.2497.
- For Landsat 8 and Landsat 9, the estimated value is 0.2304.

The final glacier identification is based on a combined condition that considers the minimum NDSI (3-2), the snow fraction (4-2), and the slope (5-2). This dual classification allows us to capture glacial data with high spectral reflectance in the near-infrared, as well as data from steep terrain with significant glacial cover.

Finally, the classification process applies a condition based on cloud fraction values (6-2), whereby pixels exceeding a value of 170 (ϴ) are classified as "no data" and are discarded from the final classification. This ensures that only pixels with ideal observations are included in the final output.

3.1. Classification variables

The classification algorithm utilizes different input variables, organized into four main categories that capture various aspects of the spectral and topographic signature of glacial surfaces.

The first category comprises six spectral bands used for glacier detection, specifically corresponding to median values from the dry season. This temporal selection is strategic, as cloud interference is minimized and the visibility of glacial surfaces is maximized during this period, allowing for better spectral discrimination between glaciers and other land cover types.

The second category includes the use of spectral indices, highlighting the temporal variations of the NDSI (Normalized Difference Snow Index) calculated for the dry season, wet season, and the minimum annual value. The use of this index is fundamental for identifying snow and glacial ice surfaces.

The third category incorporates the topographic variable of slope, derived from the SRTM Digital Elevation Model. This is crucial given that tropical glaciers are typically located in high-mountain terrain with steep topography, providing an additional geomorphological criterion for classification.

Finally, variables derived from Spectral Mixture Analysis (SMA) are included. SMA is a processing technique that decomposes the reflectance of each pixel into fractions of different pure materials, providing quantitative information about the sub-pixel composition of each area. The snow fraction is particularly relevant for distinguishing glacial surfaces from other cover types in high-mountain environments, while the cloud and shade fractions help identify and filter pixels with compromised observations.

Table 1: Bands used for classification

Type	**Name**	**Formula**	**Description**	**Reductor**[1]				
				Median	***Median dry***	***Median wet***	***Min***	***Max***
Band	*Blue*	B1 (L5 y L7) B2 (L8)	Visible spectrum Blue		X			
	Green	B2 (L5 y L7) B3 (L8)	Visible spectrum Green		X			
	Red	B3 (L5 y L7) B4 (L8)	Visible spectrum Red		X			
	NIR	B4 (L5 y L7) B5 (L8)	Near Infrared		X			
	SWIR 1	B5 (L5 y L7) B6 (L8)	Short Wave Infrared 1		X			
	SWIR 2	B7 (L5) B8 (L7) B7 (L8)	Short Wave Infrared 2		X			

Table 2: Spectral indices used for classification

Type	Name	Formula	Description	Reductor[1]				
				Median	***Median dry***	***Median wet***	***Min***	***Max***
Indices	*NDSI*	$\frac{(Green - SWIR)}{(Green + SWIR)}$	Normalized Snow Differential Index		X			

Table 3: SMA used for classification

Tipo	Nombre	Fórmula	Descripción	Reductor[1]				
				Median	***Median dry***	***Median wet***	***Min***	***Max***
Fractions	*Cloud Fraction*	SMA	Cloud fraction		X			
	Snow Fraction	SMA	Snow fraction		X			

Note 1: The reducer is based on the NDSI index, using percentiles; the 75th and 25th percentiles for Wet and Dry NDSI respectively.

3.2. Reference maps

The study area (Figure 5) was defined based on the Randolph Glacier Inventory (RGI Consortium, 2017, p. 0). Furthermore, it was visually inspected, and several missing glaciers were added.

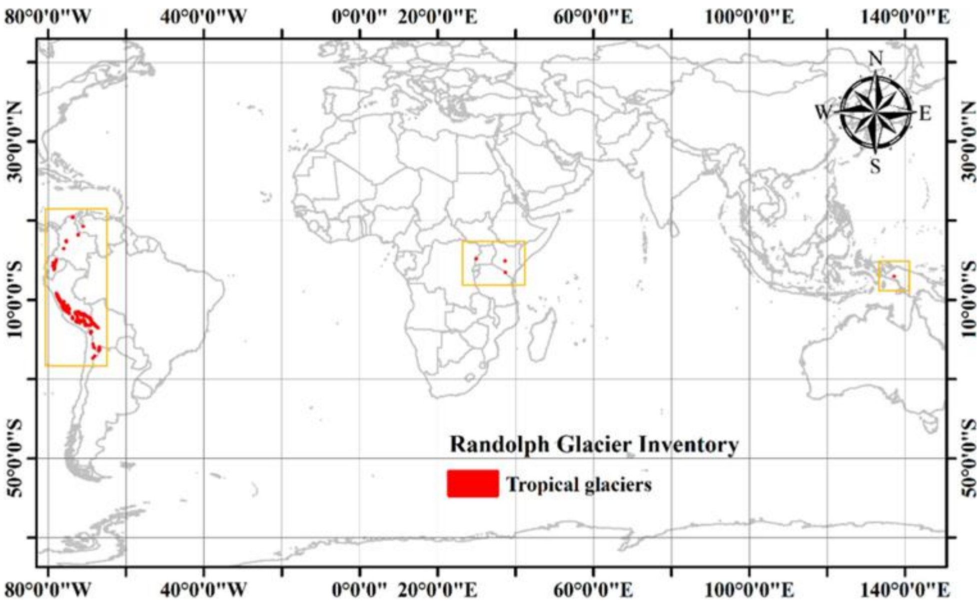

Figure 5: Distribution of tropical glaciers; source: (RGI Consortium, 2017, p. 0; Veettil & Kamp, 2019)

4. Post Classification

Given the pixel-based nature of the classification model and the processing of an extensive time series, a post-classification filter chain was implemented. This process included the sequential application of fill, temporal, spatial, frequency, and overestimation filters.

4.1. Filling Information Gaps (Gap Fill)

The filter sequence begins with a post-processing step called Gap Fill, used to fill voids in the glacier classification data. The primary objective of this filter is to correct missing or misclassified pixels within a time series. This process is fundamental for analyzing glacier time series, as it addresses one of the main challenges in remote sensing: the presence of no-data or misclassified pixels due to factors like cloud cover, topographic shadows, or limitations in satellite data capture.

The Gap Fill function implements a bidirectional approach for gap filling. This algorithm operates in two complementary phases that ensure the completeness of the analyzed time series.

In the first phase, the algorithm traverses the time series from the initial evaluation year to the final year. During this forward pass, when the algorithm encounters a no-data pixel in a specific year, it assigns the value from the immediately preceding year.

The second phase implements a reverse pass, from the final year back to the initial one. Pixels that remained as no-data after the first phase receive values from the nearest subsequent year. This bidirectional strategy ensures that data gaps persist only if a pixel lacks valid information across the entire time series.

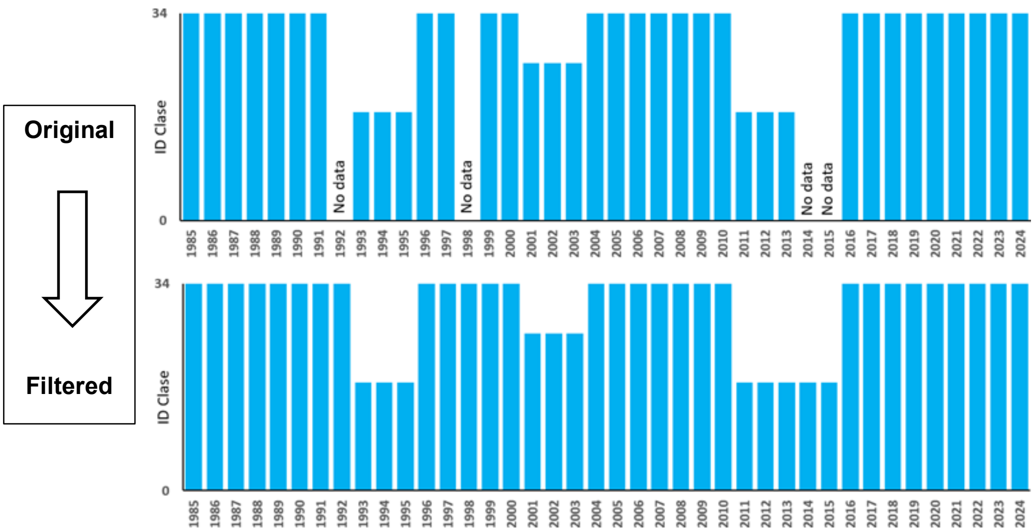

Figure 6: Gap Fill flter

4.2. Temporal Filter

The second filter implements a temporal analysis that constitutes a critical process for refining glacier classifications after the Gap Fill procedure. This filter focuses on correcting erroneous or inconsistent classifications in individual pixels throughout the time series by applying a temporal coherence logic based on multi-year analysis windows.

This filter is especially important for removing "noise" in the classifications—that is, abrupt and illogical changes in glacier cover that do not correspond to natural processes. For example, a pixel classified as "glacier" that suddenly appears as "non-glacier" for one year and then reverts to glacier likely represents a classification error rather than a real change in coverage.

The filter implements three types of temporal analysis windows: 3-year, 4-year, and 5-year windows. Each window applies specific logic to detect and correct inconsistencies in the classification.

- The 3-year window detects situations where a pixel has the same classification in the preceding and subsequent years but a different classification in the central year. This scenario indicates a classification error, so the algorithm corrects the central year to maintain temporal coherence.
- The 4-year window extends this logic to detect two consecutive years of gaps. If a pixel has the same classification for year t1 and t4, but different classifications in t2 and t3, the algorithm corrects both intermediate years. This is useful for correcting classification errors that persist for two consecutive years.
- The 5-year window handles complex cases where there are three consecutive years of erroneous classifications between two years with the same correct classification. This window is particularly useful for correcting prolonged periods of cloud cover or systematic issues in the satellite imagery.

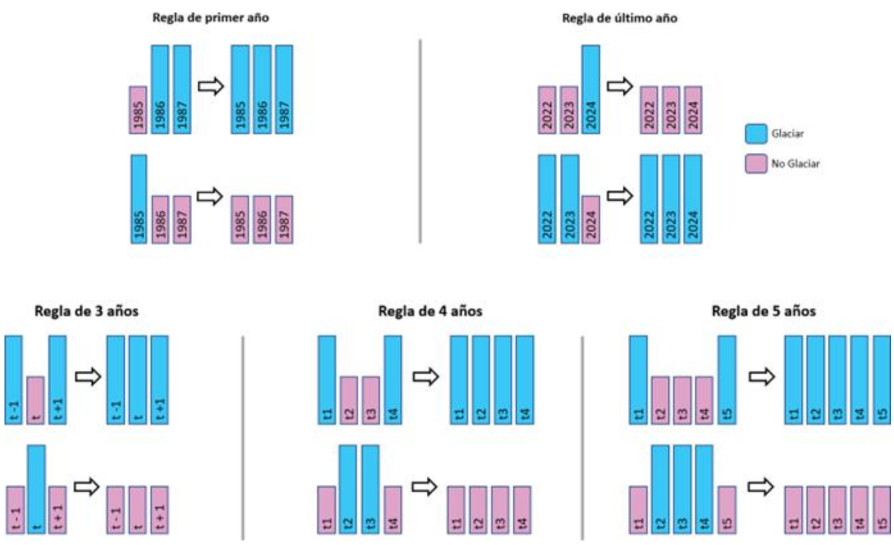

Figure 7: Temporal filter

The filter also implements specific functions to handle the endpoint years of the time series.

- For the first year, the filter checks if the two subsequent years share the same classification. If so, it adjusts the classification of the initial year to maintain coherence.

- An analogous process is applied to the final year of the series, verifying its consistency with the two preceding years.

This special treatment is necessary because these endpoint years lack a complete temporal context, requiring an adapted filtering logic. This ensures the entire time series maintains coherence from start to finish.

The order in which these rules are applied is not arbitrary. The temporal filter begins with the smaller windows, processing the endpoint years first. This stage corrects the simplest and most obvious errors before proceeding to the larger windows, thereby leaving the more complex cases for later stages. This approach maximizes the effectiveness of the filtering while minimizing the risk of overestimation.

4.3. Frequence Filter

The third filter in the processing chain is the Frequency Filter. This filter operates on the principle that stable land covers, such as glaciers, must maintain a consistent presence over time. Unlike previous filters that work with specific temporal windows, the frequency filter analyzes the entire time series to identify and consolidate the most persistent classifications.

This filter is effective at distinguishing between real changes in glacier cover and sporadic classification errors. By establishing minimum frequency thresholds, the algorithm can determine with high confidence which pixels truly represent permanent glaciers versus those that have been misclassified.

The established threshold of 70% defines the minimum percentage of years a pixel must be classified as either "glacier" or "non-glacier" to be definitively considered as such. In other words, if a pixel has been classified as a glacier in more than 70% of the years, it is permanently consolidated as a glacier. The same principle applies to the "non-glacier" class.

Figure 8: Frequence filter

4.4. Temporal Persistence Filter

The Temporal Permanence Filter complements the Frequency Filter. While the Frequency Filter evaluates the statistical persistence of classifications across the entire time series, the Temporal Permanence Filter implements a logic of irreversibility for changes in glacial cover.

This filter operates on the glaciological principle that once an area loses its glacial cover, it is extremely unlikely to become a glacier again in the short term. This premise is based on the understanding that glacier formation requires specific climatic conditions sustained over prolonged periods, which do not occur on decadal timescales (Marangunic, 2016).

Figure 9: Temporal Persistence Filter

4.5. Spatial Filter

The Spatial Filter works exclusively with the spatial dimension of the images, analyzing the coherence and connectivity of classified pixels within each individual study year.

This filter addresses a common problem in satellite image classification: the presence of isolated pixels or small pixel clusters that are classified differently from their immediate surroundings. These isolated pixels frequently represent classification errors rather than actual terrain features.

The spatial filtering process is implemented in two main stages, applied to each year in the time series.

- In the first stage, the filter uses the connectedPixelCount function to calculate how many glacier pixels are connected to each other.
- The second stage consists of grouping pixels. The filter flags all groups containing fewer than 5 connected pixels. These small groups are considered potential errors because real glaciers are rarely that small.

For glacier mapping in the Andes, where glaciers can have complex shapes due to the mountainous topography, the spatial filter is highly valuable. It helps eliminate misclassifications caused by topographic shadows, undetected small clouds, or variations in snow reflectance, while preserving the actual structure of the glaciers.

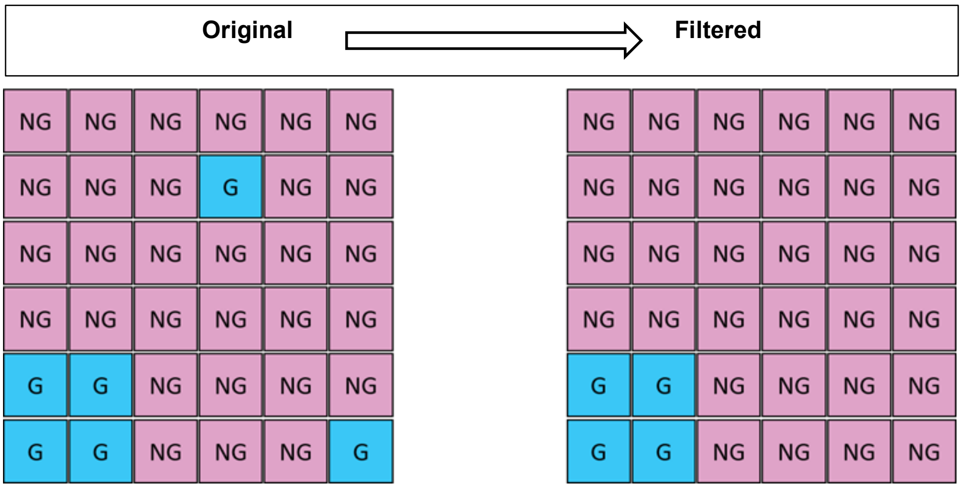

Figure 10: Spatial filter effect

4.6. Lake Masking

The masking process addresses a specific but critical problem in glacier mapping: the spectral confusion between glaciers and water bodies. This Masking Filter utilizes data from the MapBiomas Water collection to identify and remove areas where proglacial lagoons have been erroneously classified as glaciers.

This confusion is common in high-mountain environments because both glacial ice and water can exhibit similar spectral responses under certain conditions, especially when glacial sediments are present or when lakes are partially frozen.

The filter processes each year independently by integrating two complementary data sources:

- The glacier classifications, which are the result of all previously applied filters.
- The water masks, specifically identified by MapBiomas Water.

The filter implements a simple exclusion logic:

- If a pixel is classified as both glacier and water, it is reclassified as water and removed from the glacier map.
- If a pixel is classified only as glacier, it is maintained as a glacier in the final result.
- If a pixel is classified only as water, it is not considered in the final glacier map.

This process reduces the occurrence of false positives in glacier mapping, thereby eliminating a common source of error.

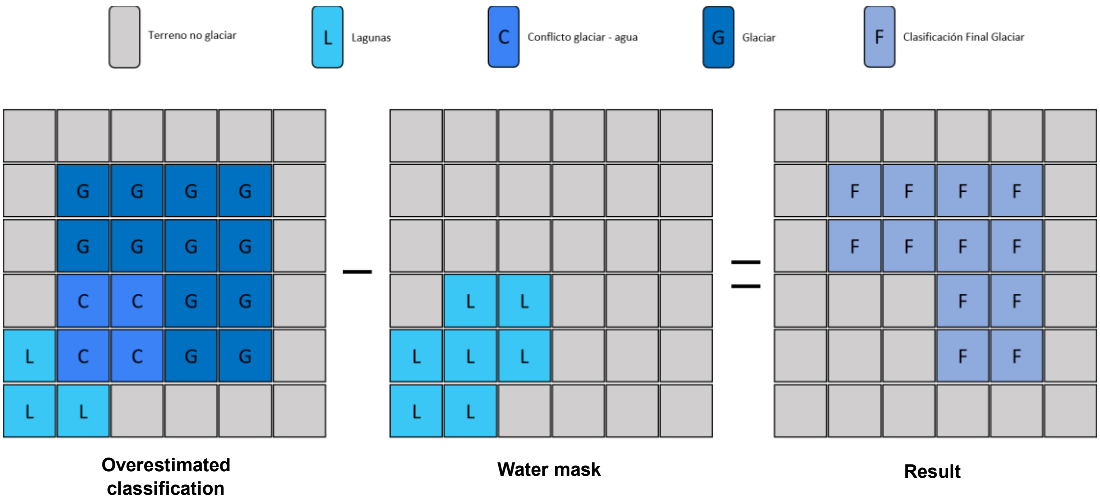

Figure 11: Lake masking sequence

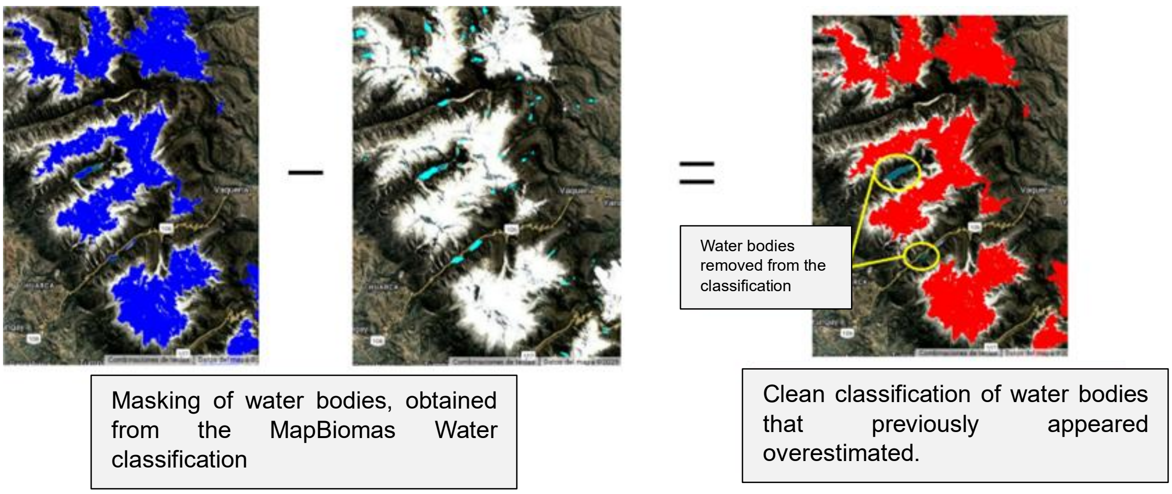

Figure 12: Lake masking effect

4.7. Integration with other MapBiomas classes

Following the application of the filter sequence in the post-classification stage, the cross-cutting themes and the general maps for each biome are integrated. This integration is governed by a set of specific hierarchical rules that assign an order of precedence to each class. The result of this stage is the production of the annual glacier cover maps for continental Ecuador.

5. Final considerations

The radiometric calibration updates implemented in Landsat Collection 1 revealed significant errors in the historical data from Landsat 4 and Landsat 5 Thematic Mapper (TM), particularly for the period 1985–1990 (Micijevic et al., 2016).

These calibration errors manifest as an overestimation of radiance and reflectance values during the approximate period of 1984–1995, with varying magnitudes depending on the spectral band. Analysis of the lifetime gain models reveals that the error is most pronounced in the early years of operation (1984–1990) and gradually attenuates towards the mid-1990s (Micijevic et al., 2016).

Cross-calibration with Landsat 7 ETM+ corrected errors of up to 2% in radiance, while integration with the Landsat 8 OLI reflectance calibration revealed even greater discrepancies (Micijevic et al., 2016).

The radiometric overestimation during the 1984–1995 period, with peak intensity between 1984–1990, has specific and differentiated consequences depending on the type of surface being mapped. For glacial and snow surfaces, the most critical bands (1, 2, and 3) show overestimations of 0.72%, 3.87%, and 4.09%, respectively. This artificially amplifies their characteristic high reflectance in the visible spectrum (Micijevic et al., 2016).

For glacial surfaces, characterized by their high reflectance in the visible spectrum, these errors would lead to an artificial overestimation of glacier area in automated classifications and would cause an inflation of multi-band spectral indices. Consequently, temporal analyses of glacial retreat that include the 1984–1995 period present an inflated initial glacier area when compared to measurements from later, correctly calibrated periods.

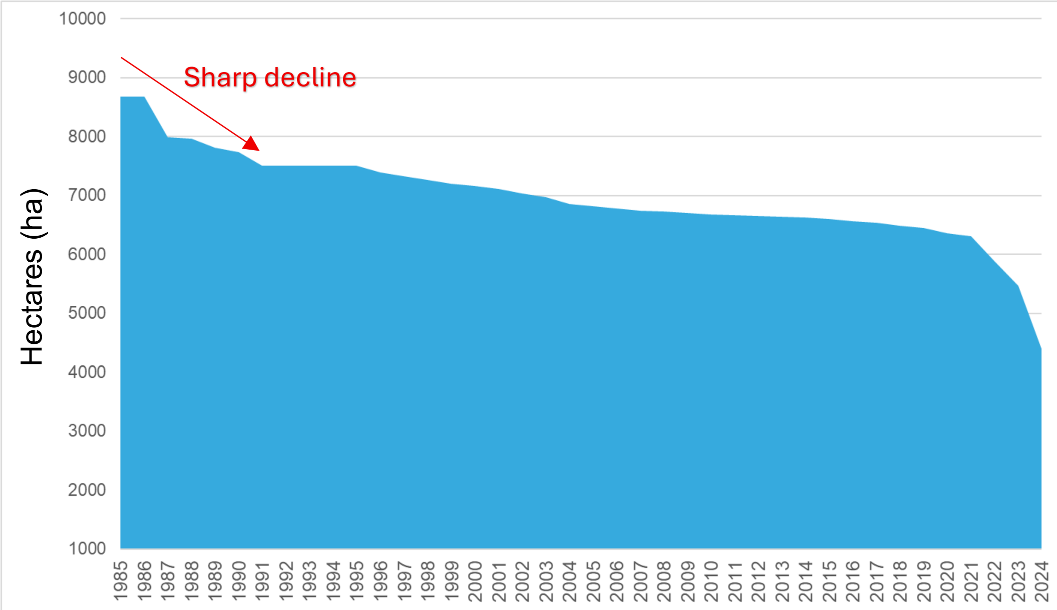

Figure 13: Glacier coverage of Ecuador assessed from 1985 to 2024

Figure 13 illustrates this phenomenon, showing a sharp decline in the recorded glacier area between 1985 and 1990. This period corresponds precisely to the maximum intensity of the calibration error. This "sharp decline" does not necessarily represent a period of real, accelerated melting but is largely an instrumental artifact resulting from the transition from overestimated data (1985) to more accurate data (toward 1990).

It is crucial that studies of glacial evolution using Landsat data from the 1984 to 1990 period interpret the apparent trends during this time with caution. Understanding this bias is essential for distinguishing between real changes and instrumental errors.

6. References

Baťka, J., Vilímek, V., Štefanová, E., Cook, S. J., & Emmer, A. (2020). Glacial Lake Outburst Floods (GLOFs) in the Cordillera Huayhuash, Peru: Historic Events and Current Susceptibility. Water, 12(10), 2664. https://doi.org/10.3390/w12102664

Huang, L., Li, Z., Zhou, J. M., & Zhang, P. (2021). An automatic method for clean glacier and nonseasonal snow area change estimation in High Mountain Asia from 1990 to 2018. Remote Sensing of Environment, 258, 112376. https://doi.org/10.1016/j.rse.2021.112376

Kaser, G., & Osmaston, H. (2002). Tropical glaciers. Cambridge Univ. Press. https://onggem.files.wordpress.com/2011/02/kaser-osmaston-2002-tropical-glaciers1.pdf

Marangunic, C. (2016). Glaciares y los ecosistemas de montaña: Las importantes tareas pendientes. Revista de Glaciares y Ecosistemas de Montaña, 1. https://doi.org/10.36580/RGEM.I1.11-19

Micijevic, E., Haque, Md. O., & Mishra, N. (2016). Radiometric calibration updates to the Landsat collection. Earth Observing Systems XXI, 9972, 99720D. https://doi.org/10.1117/12.2239426

Philip, G., & Ravindran, K. (1998). Glacial mapping using landsat thematic mapper data: A case study in parts of gangotri glacier, NW himalaya. Journal of the Indian Society of Remote Sensing, 26(1-2), 29-34. https://doi.org/10.1007/BF03007337

Racoviteanu, A. E., Paul, F., Raup, B., Khalsa, S. J. S., & Armstrong, R. (2009). Challenges and recommendations in mapping of glacier parameters from space: Results of the 2008 Global Land Ice Measurements from Space (GLIMS) workshop, Boulder, Colorado, USA. Annals of Glaciology, 50(53), 53-69. https://doi.org/10.3189/172756410790595804

RGI Consortium. (2017). Randolph Glacier Inventory – A Dataset of Global Glacier Outlines: Version 6.0. En GLIMS Technical Report. https://www.glims.org/RGI/00_rgi60_TechnicalNote.pdf

Turpo Cayo, E. Y., Borja, M. O., Espinoza-Villar, R., Moreno, N., Camargo, R., Almeida, C., Hopfgartner, K., Yarleque, C., & Souza, C. M. (2022). Mapping Three Decades of Changes in the Tropical Andean Glaciers Using Landsat Data Processed in the Earth Engine. Remote Sensing, 14(9), 1974. https://doi.org/10.3390/rs14091974

Veettil & Kamp. (2019). Global Disappearance of Tropical Mountain Glaciers: Observations, Causes, and Challenges. Geosciences, 9(5), 196. https://doi.org/10.3390/geosciences9050196