remote sensing

Article

Mapping *Polylepis* Forest Using Sentinel, PlanetScope Images, and Topographical Features with Machine Learning

Diego Pacheco-Prado [1,2,*], Esteban Bravo-López [1,3] and Luis Á. Ruiz [2]

1 Instituto de Estudios de Régimen Seccional del Ecuador (IERSE), Universidad del Azuay, Cuenca 010204, Ecuador; pbravo@uazuay.edu.ec
2 Geo-Environmental Cartography and Remote Sensing Group (CGAT), Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain; laruiz@cgf.upv.es
3 Department of Cartographic, Geodetic and Photogrammetric Engineering, Photogrammetric and Topometric Systems Research Group, Centre for Advanced Studies in Earth Sciences, Energy and Environment, University of Jaen, 23071 Jaen, Spain
* Correspondence: dpacheco@uazuay.edu.ec

Abstract: Globally, there is a significant trend in the loss of native forests, including those of the *Polylepis* genus, which are essential for soil conservation across the Andes Mountain range. These forests play a critical role in regulating water flow, promoting soil regeneration, and retaining essential nutrients and sediments, thereby contributing to the soil conservation of the region. In Ecuador, these forests are often fragmented and isolated in areas of high cloud cover, making it difficult to use remote sensing and spectral vegetation indices to detect this forest species. This study developed twelve scenarios using medium- and high-resolution satellite data, integrating datasets such as Sentinel-2 and PlanetScope (optical), Sentinel-1 (radar), and the Sigtierras project topographic data. The scenarios were categorized into two groups: SC1–SC6, combining 5 m resolution data, and SC7–SC12, combining 10 m resolution data. Additionally, each scenario was tested with two target types: multiclass (distinguishing *Polylepis* stands, native forest, Pine, Shrub vegetation, and other classes) and binary (distinguishing *Polylepis* from non-*Polylepis*). The Recursive Feature Elimination technique was employed to identify the most effective variables for each scenario. This process reduced the number of variables by selecting those with high importance according to a Random Forest model, using accuracy and Kappa values as criteria. Finally, the scenario that presented the highest reliability was SC10 (Sentinel-2 and Topography) with a pixel size of 10 m in a multiclass target, achieving an accuracy of 0.91 and a Kappa coefficient of 0.80. For the *Polylepis* class, the User Accuracy and Producer Accuracy were 0.90 and 0.89, respectively. The findings confirm that, despite the limited area of the *Polylepis* stands, integrating topographic and spectral variables at a 10 m pixel resolution improves detection accuracy.

Keywords: Andes Mountains; Google Earth Engine; NICFI; random forest; RFE

check for updates

Citation: Pacheco-Prado, D.; Bravo-López, E.; Ruiz, L.Á. Mapping *Polylepis* Forest Using Sentinel, PlanetScope Images, and Topographical Features with Machine Learning. *Remote Sens.* **2024**, *16*, 4271. https://doi.org/10.3390/rs16224271

Academic Editors: Ali Nadir Arslan, Jose Manuel Álvarez-Martínez and Katarzyna Dabrowska-Zielinska

Received: 8 October 2024
Revised: 12 November 2024
Accepted: 14 November 2024
Published: 16 November 2024

1. Introduction

In the South American Andes, the *Polylepis* forests hold significant conservation importance due to their roles in regulating water flow, facilitating soil regeneration, and retaining essential nutrients and sediments [1–4]; however, they are some of the most threatened ecosystems [5] mainly due to anthropogenic activities such as deforestation [1,2], agriculture [6,7], large-scale cattle ranching, and demographic pressure [5,8]. Therefore, detecting the location of *Polylepis* stands is crucial for maintaining these ecosystems, particularly in Ecuador, which experiences high deforestation rates and significant forest loss along the Andean slopes [6]. For instance, the remnants of native forests have been reduced from approximately 68% to 56% between 1990 and 2018 [9].

Polylepis is generally found in protected areas (near lagoons, in rocky areas, canyons, and stream banks) of the Andes Mountains [10,11] in Venezuela, Colombia, Ecuador, Peru, Bolivia, Chile, and Argentina, in altitudinal ranges from 1800 to 5200 m above sea level (m.a.s.l.), in isolated stands with limited surface area and difficult accessibility [10]. However, certain factors that influence the distribution of *Polylepis* trees are still not known in depth [12].

On the other hand, remote sensing techniques have been used to identify, characterize, and assess forest species in biodiversity conservation and ecosystem management. For example, the review by Kivinen et al. of *Populus tremula* L., an ecologically important species in boreal forests, identified the importance of integrating ecological knowledge with remotely sensed information to understand the biodiversity patterns of this species, as well as its suitability for mapping [13]. In other studies, Schwager and Berg used remote sensing to improve Species Distribution Models (SDMs) to determine the distribution of alpine plant species [14]. Similarly, but in the Qilian Mountains (China), Ma et al. classified four tree species: *Sabina przewalskii*, *Picea crassifolia*, *Betula*, and *Populus* using satellite imagery (specifically Sentinel-2 imagery) [15]. Remote sensing has also made it possible to assess spatiotemporal changes in forest cover, biomass, plant species composition, and health, thus improving the ability to estimate biomass at the level of individual trees [16]; classify forests into species groups [17]; assessing the health status of urban forests [18]; understanding succession patterns in tropical forests to monitor their biodiversity [19]; or analyzing forest recovery under diverse climatic conditions [20].

Using specific satellite images also makes identifying and characterizing species possible, as evidenced in the works cited above. Another important aspect refers to the Sentinel mission images. Sentinel is a satellite constellation part of the Copernicus program of the European Space Agency (ESA), providing free and medium-resolution satellite imagery. In this context, two of the missions are Sentinel-1 (S1) and Sentinel-2 (S2). The first (S1) has a Synthetic Aperture Radar (SAR) sensor that generates its energy wave and measures backscattered radiation from the ground [21] and can penetrate cloud cover and obtain data in different atmospheric conditions [22]. In addition, S1 mission data are used in the construction of forest masks [23], in the assessment of physical characteristics and properties such as the roughness, water content (moisture), and structural geometry of forest species [24], and in the detection of forest degradation and regeneration [25]. However, it can have substantial distortions due to the effect of topography [23]. S2 optical images are used in the detection and monitoring of land cover and other applications through the MultiSpectral Instrument (MSI). Thanks to the operation of two similar satellites, S2A and S2B, the image acquisition of the same area is achieved every five days. Its usefulness is reflected in the different applications of these images for the detection of invasive species [26], the identification and classification of tree species [27], or the use of multitemporal data for the characterization of regenerating forest stands [28]. To enhance the discrimination of forest species, various studies have demonstrated the effectiveness of combining S1 and S2 data [21,24], as well as their integration with topographic information [29], in Cloud Computing platforms such as Google Earth Engine (GEE) [30–32]. On the other hand, there is currently spectral information with better spatial resolution available through the PlanetScope platform, acquiring images with 3 m/pixel spatial resolution and capturing data in the red (R), green (G), blue (B), and near-infrared (NIR) channels. However, the open-access data provided by Norway's International Climate and Forests Initiative (NICFI) program, in collaboration with Planet Labs of the PlanetScope data, are available in semiannual or monthly mosaics with a resolution of 4.77 m per pixel. Some studies, such as the one from Vizzari (2022), use these data in the GEE environment to compare ten land use and land cover (LULC) classification approaches, highlighting the advantages of adopting an object-based (OB) approach with PlanetScope (PL) data and integrating PL with Sentinel-2 and Sentinel-1 data. These studies incorporate Gray-Level Co-occurrence Matrix texture variables to aid classification [33].

Although remote sensing is widely applied in identifying, monitoring, and characterizing forest species, its implementation presents important challenges in the study of the genus *Polylepis*. Some studies are focused on this species, such as the one conducted by

López et al., who evaluated the distribution and cover of a *Polylepis* species (*P. tarapacana*) and its relationship with topographic, environmental, and climatic factors in the Andean zone of northern Argentina, identifying that elevation and slope are significant factors in the distribution of this species within the study sector [11]. Caballero-Villalobos et al. [12] evaluated the possible distribution of *Polylepis quadrijuga* in the Colombian Andes due to its ecological importance, using different future temporal scenarios, demonstrating that climate change could generate a considerable loss of suitable habitat for this species. Rojas et al. evaluated the unique climatic habitat conditions of *Polylepis Tarapacana* in the Chilean Andes. They highlighted the urgent need to develop studies that analyze adaptive strategies for this species to withstand climate changes at high altitudes [34]. It is worth noting that studies employing pixel and object classification methods [35,36] within the mountainous forests of the Andean region have major limitations [22], mainly due to the topographic complexity, shadows, and illuminated areas [35], the limited availability of cloud-free satellite images [8,37], and the challenge of finding completely homogeneous pixels [22]. However, there are approximations through the detection of native forests with the combination of optical and radar features [38].

Combining remote sensing data with image processing and classification methods is useful to analyze and identify forest entities [39]. Several works present the feasibility of the evaluation of forest ecosystems through Machine Learning (ML) algorithms such as Random Forest (RF) [22,30,35,36,38,40–42], Support Vector Machines [22,35,36,43], Gradient Boosting [44], Naive Bayes [45], being RF one of the most used, especially in the discrimination and identification of mountain forests and shrublands [22]. Additionally, some researchers used hyperspectral images obtained from Unmanned Aerial Vehicles (UAVs) for the identification of *Polylepis* [10] and other forest species [46–48].

Despite this progress, several challenges remain in identifying *Polylepis* using remote sensing. These challenges include (1) The small size of *Polylepis* stands limits their detection with freely available satellite data. (2) Accessibility to these forests is complicated, making it challenging to obtain field validation data. (3) The spectral signature of *Polylepis* is similar to those of other shrub vegetation in the study zone.

Given this background, the objective of this study is to (a) Compare classification scenarios at spatial resolutions of 5 and 10 m and evaluate if there is any difference in the efficiency of *Polylepis* detection based on the spatial resolution of the input data; (b) evaluate the combination of optical, radar, and topographic data for *Polylepis* stands identification; and (c) identify the minimum variables needed to classify *Polylepis* stands using a pixel-based approach.

2. Materials and Methods

2.1. Study Zone

The study area is located in the Andes Mountains in southern Ecuador (Figure 1), with an approximate area of 137,211 ha. It has an altitude range from 918 to 4534 m.a.s.l. The predominant land cover is paramo, an ecosystem where *Polylepis* species are present [49], occupying approximately 45% of the area. Vegetation classes such as native forests, *Polylepis* stands, forest plantations, and shrublands (chaparro) account for approximately 24% of the area, with *Polylepis* being one of the least represented, covering less than 1%. *Polylepis* stands are believed to be "possible remnants of vegetation that once covered larger areas" [50]. The remaining area corresponds to other classes, such as water bodies and bare soil.

Although a significant portion of the study area falls within the protected boundaries of Cajas National Park (CNP), it remains vulnerable to deforestation and associated threats [6].

2.2. Polylepis Stands

Polylepis stands are distributed along the western and eastern sierras [49] and in Ecuador. According to the Global Biodiversity Information Facility (GBIF) database https://www.gbif.org/es/ (accessed on 25 July 2022), these stands can be found above 1067 m.a.s.l. Their height ranges from 8 to 10 m, with twisted branches covered with moss [49]. Ecuador is the eighth country with the greatest variety of tree species in the world [51]. Of the genus *Polylepis* (identification target), seven native species can be found: *P. incana*, *P. lanuginosa*, *P. microphylla*, *P. pauta*, *P. reticulata*, *P. sericea*, and *P. weberbaueri*. In the study area (southern Ecuador), four of these species are found: *P. reticulata* (Figure 2), *P. lanuginosa*, *P. weberbaueri*, and *P. incana* [49].

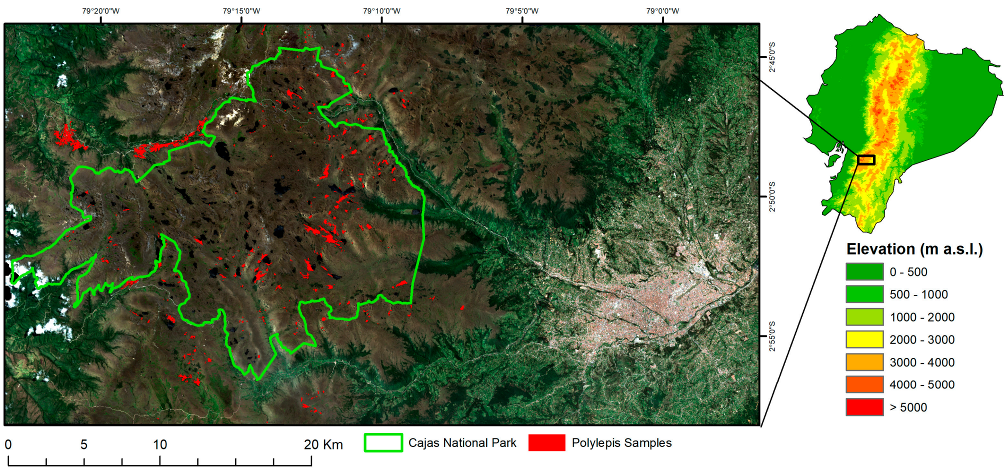

Figure 1. Sentinel-2A image mosaic covering tiles 17MPS, 17MPT, 17MQS, and 17MQT (24 August 2020) of the study area and location of *Polylepis* stands.

(**a**) (**b**)

Figure 2. *Polylepis reticulata* tree (**a**) and stand (**b**) in Toreadora lagoon zone.

2.3. Workflow Description

With the vector and satellite data available, a methodology was proposed that can be explained in four stages (Figure 3). The first stage consisted of searching for information on the location of *Polylepis* stands in the study area and integrating it with a land use and land cover (LULC) map, which generated a base map from which the target classes of the classification model were determined (green section). In the second stage, satellite data were selected and prepared through the GEE platform, including S2 and PlanetScope spectral data (bands and indices) and textures of the NIR band of PlanetScope image, S1 radar data, and the Digital Elevation Model (DEM) obtained from the Sigtierras project [52] (yellow section). In the third stage, the scenarios to be evaluated were defined, considering the spatial resolution and the combination of different types of variables (orange section). The final stage consisted of comparing different scenarios and variables to determine the combination that produces the highest reliability in detecting *Polylepis* stands. Metrics from each scenario were evaluated, and the best-performing scenario was selected. Using this scenario, a *Polylepis* map was generated and validated by forestry specialists in the region (blue section).

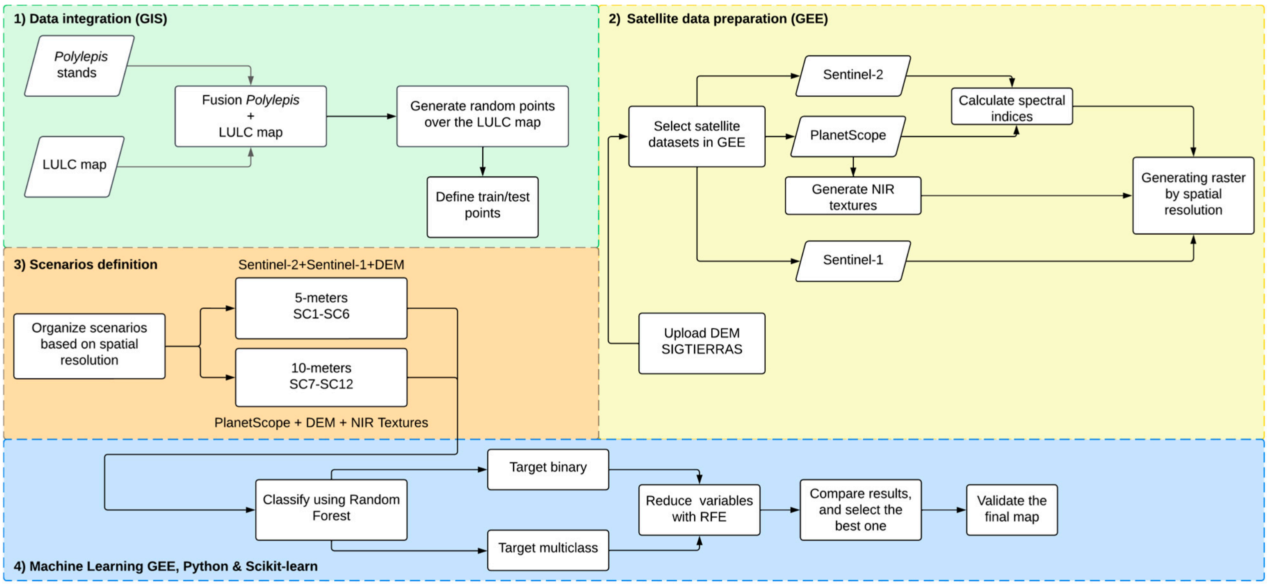

Figure 3. Workflow diagram.

2.3.1. Data Integration

The LULC map of the study area was obtained from the Azuay province at a scale of 1:5000 [53], which was constructed from the aerial orthophotography of the SIGTIERRAS project of Ecuador [52]. The *Polylepis* occurrence data for the study area were sourced from georeferenced records available in online repositories, including the Global Biodiversity Information Facility (GBIF) and Herbario Azuay, collected in 2016. These initial records were validated through field verification conducted between July 2016 and October 2017. The final dataset comprises 244 presence records of the *Polylepis* genus distributed as follows: 55 records for *P. incana*, 86 for *P. lanuginosa*, 50 for *P. reticulata*, and 53 for *P. weberbaueri*. Of the total dataset, 73% represents field-verified observations, while the remaining 27% corresponds to botanical collection data or herbarium records.

Based on this LULC map, we generated random points in two targets, multiclass and binary, for the classification processes (Table 1). In the binary target, the No *Polylepis* class groups classes such as native forests, pine, shrub vegetation, and non-forest classes. Finally, seventy percent of the points were allocated for training, while the remaining thirty percent were reserved for validation.

Table 1. Forest classes and points (pixels) number for training and validation in multiclass and binary targets.

		Multiclass Target					Binary Target		
Class Name	Code	Train	Validation	Total	Class Name	Code	Train	Validation	Total
Polylepis	1	2362	974	3336	*Polylepis*	1	2362	974	3336
Native forest	2	303	150	453	No *Polylepis*	2	11,732	5113	16,845
Pine	3	93	35	128					
Shrub vegetation	4	1674	700	2374					
Other classes	5	9662	4228	13,890					

2.3.2. Satellite Data Preparation

Satellite data acquisition (Table 2) and the generation of satellite information indices (Table 3) were conducted using the GEE tool. This Cloud Computing approach enabled multiple tasks to be performed on a high volume of data [54], facilitating the workflow. The data used were as follows:

Table 2. Multispectral, SAR, and topographic datasets used in the study.

Dataset	Temporality	Spatial Resolution/Bands	Reference
Sentinel-2A MSI	Mean value August 2020	10 m (TCI_R, TCI_G, TCI_B, B2, B3, B4, B8) 20 m (B5, B6, B7, B8A, B11, B12) 60 m (B1, B9)	[55]
Sentinel-1 SAR GRD	Mean value August 2020	10 m (VH, VV)	[23,56]
PlanetScope normalized analytic	Mosaic June 2020–August 2020	4.77 m (R, G, B, NIR)	[33]
Topographic data	Photogrammetric fly 2010	3 m (Elevation, Slope, Aspect)	[56]

Table 3. Vegetation indices and formulas used in this study.

Indices	Formula	Reference
EVI Enhanced Vegetation Index	$2.5 * \frac{NIR-Red}{(NIR+6Red-7.5Blue)+1}$	[57]
SIPI Structure Insensitive Pigment Index	$\frac{NIR-Aerosols}{NIR-Red}$	[58]
NDWI Normalized Difference Water Index	$\frac{Green-NIR}{Green+NIR}$	[56]
MSI Simple Ratio 1600/820 Moisture Stress	$\frac{SWIR\ 1}{NIR}$	[41]
ARI Anthocyanin Reflectance Index	$\frac{1}{Green} - \frac{1}{Red\ Edge\ 1}$	[58]
CCCI Canopy Chlorophyll Content Index	$\frac{\frac{NIR\ -\ Red\ Edge\ 1}{NIR\ +\ Red\ Edge\ 1}}{\frac{NIR\ -\ Red}{NIR\ +\ Red}}$	[58]
MSR Modified Simple Ratio 670/800	$\frac{NIR-Blue}{Red+Blue}$	[58]
GLI Green Leaf Index	$\frac{2Green-Red\ Edge\ 1-Blue}{2Green+Red\ Edge\ 1+Blue}$	[58]
NDVI Normalized Difference Vegetation Index	$\frac{NIR-Red}{NIR+Red}$	[33,41,56,57]
NGRDI Normalized Green Red Difference Index	$\frac{Green-Red\ Edge\ 1}{Green+Red\ Edge\ 1}$	[58]

(a) Sentinel-2 optical data (level 2A), radiometrically corrected to ground reflectance level, were used. There data correspond to the mosaic of tiles 17MPS, 17MPT, 17MQS, and 17MQT from 24 August 2020 to compose an image with the minimum presence of clouds. In GEE, the COPERNICUS/S2_SR_HARMONIZED repository was used in which the values of the scenes were harmonized so that the most recent data (after 25 January 2020) are in the same range as the oldest scenes [55].

(b) Synthetic Aperture Radar (SAR) data in Interferometric Wide (IW) mode and processed in Ground Range Detected (GRD) format were obtained from the Sentinel-1 sensor. This sensor offers two polarities: VV (vertical transmit/vertical receive) and VH (vertical transmit/horizontal receive).

(c) Norway's International Climate and Forests Initiative (NICFI) provides free access to high-resolution PlanetScope imagery in Google Earth Engine (GEE). This dataset includes the red, green, blue, and near-infrared (NIR) bands, available as a mosaic spanning from June 2020 to August 2020, and resampled from 4.77 m to 5 m of pixel size. The textural features obtained from the NIR band were extracted via Gray-Level Co-occurrence Matrix (GLCM). The features generated include Max Corr. Coefficient (N_maxcorr), Angular Second Moment (N_asm), Entropy (N_ent), Cluster Prominence (N_prom), Difference Entropy (N_dent), Sum Entropy (N_sent), Variance (N_var), Inertia (N_inertia), Information Measure of Corr. 1 (N_imcorr1), Information Measure of Corr. 2 (N_imcorr2), Contrast (N_contrast), Difference Variance (N_dvar), Sum Variance (N_svar), Correlation (N_corr), Dissimilarity (N_diss), Shade (N_shade), and Inverse Difference Moment (N_idm).

(d) Topography data were obtained from the Ecuador Sigtierras project, which provides data distributed with a pixel size of 3 m. This dataset was used to determine the terrain topography and derive slope and orientation, resampled to a pixel size of

5 and 10 m using bilinear interpolation. Due to the varying spatial resolutions of the satellite data, the data were grouped into categories of 5 m and 10 m resolutions.

2.3.3. Scenarios Definition

After defining the datasets to be used, we evaluated different scenarios while maintaining spatial resolutions of 5 m (RF1-RF6) and 10 m (RF7-RF12). This analysis aimed to determine whether spatial resolution influenced the outcomes, as summarized in Table 4. The datasets are represented as follows: "TOPO" refers to Digital Elevation Model data from Sigtierras, including slope and aspect derivatives. "PS" represents PlanetScope imagery and indexes, while "GLMC" denotes textures derived from the near-infrared band of PS. "S2" represents Sentinel-2 data and indexes, and "S1" represents Sentinel-1 data. In the proposed scenarios, the multiclass or binary target approach will also be evaluated to distinguish *Polylepis* from other shrub classes.

Table 4. Scenario definition based on dataset combination.

Scenario	Resolution	Dataset Combination	Variables Number
SC1	5 m	TOPO	3
SC2		PS	7
SC3		GLCM NIR	18
SC4		TOPO + PS	10
SC5		TOPO + GLCM NIR	21
SC6		PS + TOPO + NIR GLCM	28
SC7	10 m	S1	2
SC8		S2	24
SC9		S1 + TOPO	5
SC10		S2 + TOPO	27
SC11		S2 + S1	26
SC12		S2 + S1 + TOPO	29

Variable Reduction

GEE provides a variety of supervised classification algorithms suitable for satellite image classification. This study employed the Random Forest (RF) algorithm, implemented in GEE as smileRandomForest (SRF). RF is a nonparametric statistical method of decision tree assembly widely used in land cover classification from remote sensing data [21]. Each tree is trained with different data so that by combining the results, some errors are compensated by others. This algorithm is used by several authors in forestry domains based on pixel classification of satellite images [22,45,59–61] and with object approach as in [36,40]. For this study, the number of trees was set to 50, while other parameters were left at their default values to facilitate scenario reliability comparison.

From the former, the Recursive Feature Elimination (RFE) technique has been selected for its ability to eliminate noise and irrelevant features, using a Machine Learning algorithm to rank features by importance, discard the least important features, and re-fit the model to improve the performance of the classification model [62]. In GEE, the SRF algorithm was employed to assess the importance of variables in each classification scenario. The variable reduction with RFE was performed by being the first iteration run with all variables, and, in the following runs, the features of lesser importance were removed. After eliminating the least significant variables, the classification process was run again, and the proposed metrics were evaluated.

Scenarios Comparation

The classification results with all variables were evaluated in a general way, using the Global Accuracy, which refers to the ratio between correctly classified pixels and all verification pixels (the main diagonal of the confusion matrix), and the Kappa coefficient to evaluate the classification performance compared to the expected random assignment [63]. Other metrics used focused on evaluating the classification performance of the *Polylepis* class, using for this purpose the User Accuracy (UA), which refers to the ratio between the total number of correctly classified samples for a particular class and the total number of reference samples classified in that same class; and the Producer Accuracy (PA), which shows the ratio between the accurately classified reference samples of a particular class and the total number of reference samples of that class [63], this in GEE.

Finally, the F-score is a harmonic mean between Precision and Recall [33], which allows the individual evaluation of class behavior. This process will be carried out on the most reliable scenarios using scikit-learn library version 1.5.2 and Python version 3.10.12.

3. Results

3.1. Variable Reduction Using RFE and Comparison of Classification Metrics for Different Scenarios

The reliability values obtained from the various classification scenarios (GEE) are compared in Table 5. This table reports the total number of variables and the variables selected by the RFE process, considering their importance in the RF algorithm. The initial variable set (N column) is significantly reduced in most scenarios through the Recursive Feature Elimination (RFE) process, underscoring its effectiveness in filtering out less informative features. For example, in SC6, the original 28 variables are pared down to 11 post-RFE (Figure 4). Conversely, RFE retains a smaller subset of features in scenarios that start with fewer variables (e.g., SC1 and SC7), suggesting that these scenarios either have inherently simpler data structures or initially contain fewer irrelevant features.

Table 5. *Polylepis* classification scenarios: spatial resolution of input features, number of features produced and selected, Accuracy (Acc), and Kappa coefficient after the RFE reduction process for binary and multiclass targets. UA and PA refer to the *Polylepis* class.

				Multiclass Target					Binary Target				
Name	**Resolution**	**Datasets**	**N ***	**RFE**	**Acc**	**Kappa**	**UA**	**PA**	**RFE**	**Acc**	**Kappa**	**UA**	**PA**
SC1	5 m	TOPO	3	3	0.73	0.37	0.46	0.30	3	0.83	0.26	0.45	0.29
SC2		PS	7	7	0.86	0.69	0.75	0.79	7	0.93	0.72	0.80	0.74
SC3		GLCM NIR	18	5	0.68	0.12	0.36	0.12	18	0.84	0.08	0.42	0.07
SC4		TOPO + PS	10	9	0.90	0.78	0.87	0.89	9	0.96	0.86	0.89	0.87
SC5		TOPO + GLCM NIR	21	7	0.76	0.41	0.64	0.25	6	0.85	0.29	0.59	0.25
SC6		PS + TOPO + NIR GLCM	28	11	0.90	0.78	0.88	0.88	7	0.96	0.86	0.90	0.87
SC7	10 m	S1	2	2	0.68	0.24	0.29	0.20	2	0.82	0.09	0.31	0.12
SC8		S2	24	8	0.88	0.74	0.85	0.85	8	0.95	0.81	0.87	0.81
SC9		S1 + TOPO	5	5	0.82	0.60	0.74	0.66	4	0.88	0.53	0.69	0.54
SC10		S2 + TOPO	27	11	0.91	0.80	0.90	0.89	7	0.97	0.87	0.90	0.88
SC11		S2 + S1	26	13	0.89	0.76	0.85	0.86	10	0.95	0.82	0.88	0.82
SC12		S2 + S1 + TOPO	29	7	0.91	0.80	0.89	0.90	9	0.96	0.86	0.90	0.87

* The column N represents the initial number of variables for each scenery. This is reduced by the RFE process and presents the new number of variables in column RFE.

The scenarios that produced the best accuracy and Kappa values are SC4 (0.90 and 0.78) for the 5 m datasets. In the 10 m datasets for the multiclass approach, SC10 (0.91 and 0.80) and SC12 (0.91 and 0.80) were the best. For the binary target, SC6 (0.96 and 0.86) and SC10 (0.96 and 0.87) were the best. On the other hand, the scenarios that evaluated the datasets without multispectral information (SC1, SC3, SC7, and SC9) had lower reliability compared to the other scenarios that included these types of variables.

In scenarios using 10 m resolution data (SC7–SC12), the models exhibit strong predictive performance across both multiclass and binary classification tasks, with particularly notable gains when Sentinel-2 imagery is combined with topographic data (TOPO). Meanwhile, in scenarios employing 5 m resolution data, the integration of PlanetScope imagery with additional layers, such as topographic or texture-based features (e.g., SC4, SC5, and SC6), consistently improves classification accuracy. These findings imply that the combined use of high-resolution data and a diverse set of feature types can mitigate the constraints imposed by the coarser resolution of certain datasets, thereby enhancing overall model performance.

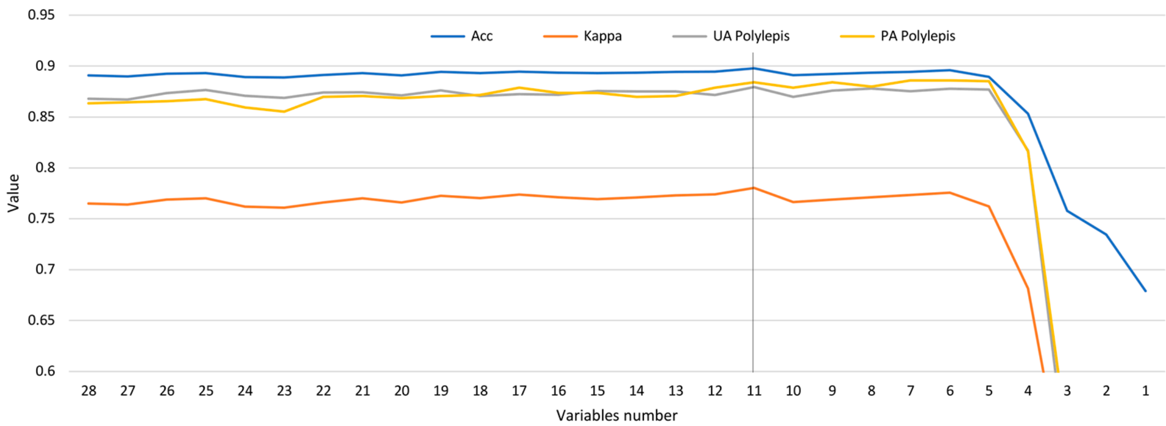

Figure 4. Behavior of Accuracy (blue line), Kappa (orange line), UA (gray line), and PA (yellow line) metrics during the RFE process in the SC6 scenario. Peak values are obtained with 11 variables.

An examination of the F1 scores (Table 6) across various land cover classes in optimal scenarios (SC4, SC6, SC10, and SC12) reveals a consistent trend in the multiclass classification. For instance, the *Polylepis* class maintains high accuracy across all configurations, ranging from 0.88 to 0.90, whereas the Pine class exhibits substantial variability, with accuracy scores ranging from as low as 0.00 in SC6 to 0.48 in SC10. In the binary classification setting, which differentiates between *Polylepis* and No Polylepis, both classes achieve robust accuracies, with scores between 0.88 and 0.98. Lower performance is observed in classes with limited sample sizes, such as Pine and Native Forest, potentially due to restricted data representation or class overlap, which may challenge the model's discrimination capacity.

Table 6. F1-score comparison by best scenarios.

Target	Class	Pixels Number	SC4	SC6	SC10	SC12
Multiclass	*Polylepis*	974	0.88	0.88	0.90	0.89
	Native forest	150	0.49	0.45	0.37	0.25
	Pine	35	0.06	0.00	0.48	0.37
	Shrub vegetation	700	0.72	0.73	0.73	0.74
	Other classes	4228	0.94	0.95	0.95	0.95
Binary	*Polylepis*	974	0.88	0.88	0.89	0.89
	No *Polylepis*	5113	0.98	0.98	0.98	0.98

In the best scenarios, several variables are repeated multiple times, indicating their importance across different scenarios. This is particularly true for topographic variables such as elevation, slope, and aspect (Figure 5). In the multispectral domain, the indices PS_NDVI (5 m), CCCI, MSR, and MSI from S2 (10 m) appeared in various classification scenarios; these indices have in common the band 8 of S2 in the range of 835.1 nm.

3.2. Comparison of the Maps of the Best Scenarios

Figure 6 presents a detailed example of the supervised classification results in a subarea of the Toreadora Lagoon, which includes a variety of shrub classes. In this region, pixels classified as *Polylepis* are shown in green. The scenario SC4, with a multiclass target and a spatial resolution of 5 m, also performs well with a high accuracy of 0.90 and Kappa of 0.78, indicating that different resolutions can yield similar results depending on the dataset combination. However, the evaluation of known areas reveals high separability between

Polylepis and other vegetation classes in the 10 m SC10 and SC12 multiclass scenarios. In the multiclass target, red polygons indicate a Pine plantation. Although the Accuracy and Kappa metrics favor the SC4 scenario in the binary classification, the UA and PA for *Polylepis*, when evaluated together, were better in the SC10 multiclass scenario. This is reflected in an improved distinction of the *Polylepis* class from other shrubs, as shown in Figure 6.

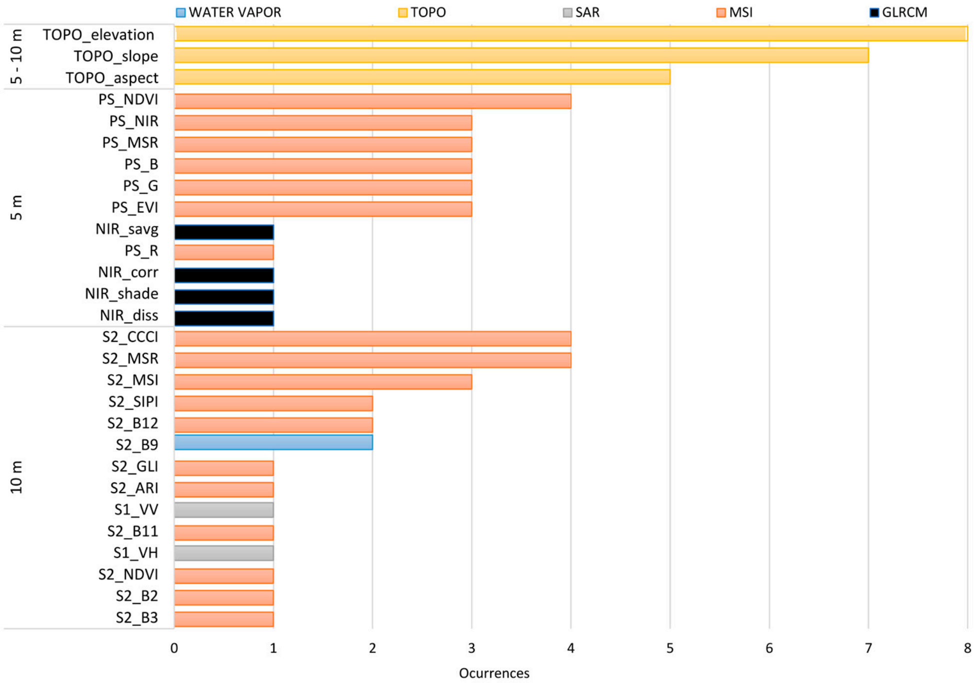

Figure 5. Variables that were repeated (occurrences) in the best eight scenarios. S2 (Sentinel-2), S1 (Sentinel-1), TOPO (Topography), and PS (PlanetScope).

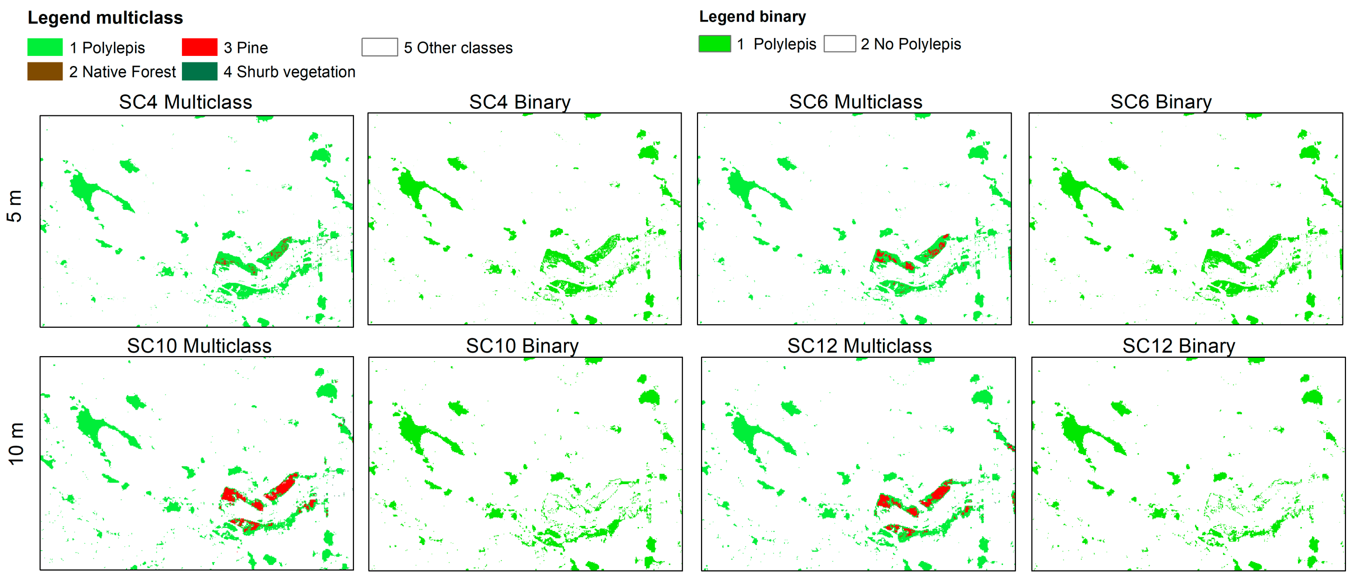

Figure 6. Detailed view of the eight best classification scenarios in the Toreadora Lagoon region.

The final scenario selected was SC10 with a multiclass target (Figure 7), showing *Polylepis* located in the paramo zone. This was achieved using variables MSI, B9, B2, MSR, CCCI, B12, ARI, and NDVI of S2 in conjunction with aspect, slope, and elevation. The variable importance analysis (Figure 8) revealed that elevation and slope were the most significant factors, while the selected spectral indices and S2 bands had a similar contribution. In this map, another known area is Llaviuco Lagoon, where a clear distinction can be observed between *Polylepis* stands and shrub vegetation despite both classes presenting similar spectral characteristics.

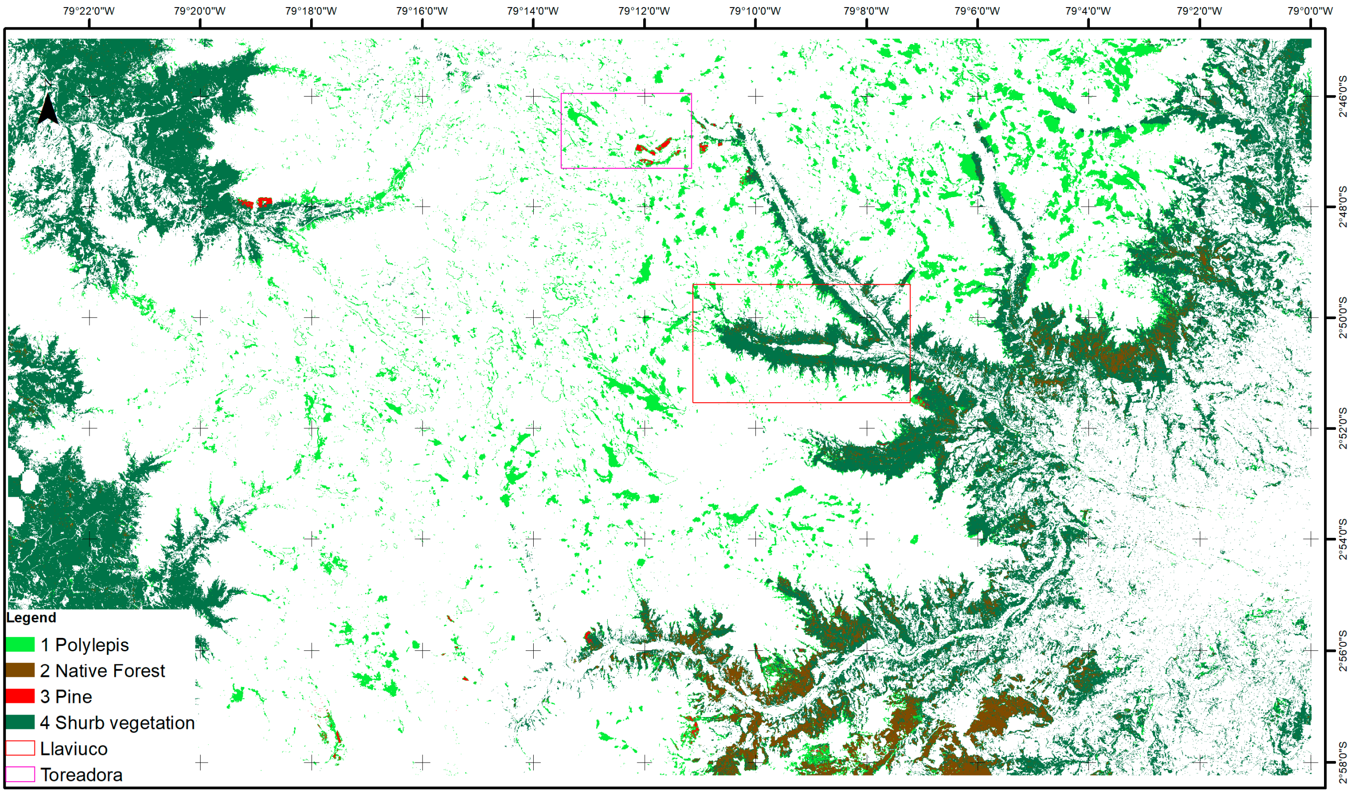

Figure 7. SC10 multiclass map was selected as the best scenario. Known regions such as Toreadora Lagoon (purple polygon) and Llaviuco Lagoon (red polygon) helped to visually interpret the quality of the classification models. Other land uses and coverages in the area are shown in white color.

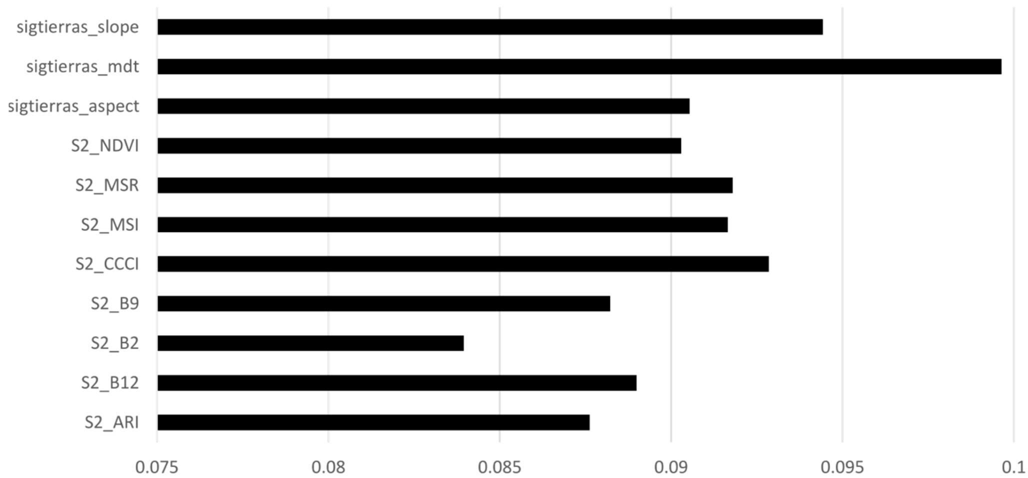

Figure 8. Feature importance of SC10 multiclass scenario.

The reports and confusion matrix of each scenario are available at https://gis.uazuay.edu.ec/papers/polylepis/ (accessed on 24 September 2024).

The map in Figure 7 presents the spatial distribution of land cover classes in the study area, focusing on the classification of *Polylepis* stands, native forest, pine plantations, shrub vegetation, and other vegetation types. *Polylepis* stands are highlighted in light green, native forests in brown, pine plantations in red, shrub vegetation in dark green, and other classes in gray. The mapped area includes two highlighted zones outlined in red and magenta, labeled as "Llaviuco" and "Toreadora", respectively, indicating specific regions of interest for *Polylepis* detection.

The spatial extent shows that *Polylepis* stands are dispersed in specific regions, mostly fragmented and interspersed with other vegetation classes. Native forest areas appear more contiguous, primarily concentrated in the lower central part of the map, while pine plantations are present in relatively small clusters, mostly located near the Llaviuco region. Shrub vegetation and other classes occupy larger contiguous areas, particularly along the central and southern regions of the map, suggesting a gradient of vegetation types influenced by altitude and possibly other topographic variables.

This map allows for a visual evaluation of classification accuracy and the spatial distribution of *Polylepis* stands within a diverse landscape matrix, providing insights into potential ecological factors affecting *Polylepis* distribution.

4. Discussion

In this work, we compared twelve scenarios resulting from combinations of Sentinel-2 (S2), Sentinel-1 (S1), PlanetScope (PS), texture GLCM for NIR PS band, and topographic data (elevation, slope, and aspect), using two target types (multiclass or binary) for the classification of *Polylepis* stands. Among all scenarios, four (SC4, SC6, SC10, and SC12) demonstrated the highest reliability in Accuracy and Kappa. Although in several scenarios, the metrics did not show significant differences, selecting known review areas allowed for a better detection of the best scenario by *Polylepis* discrimination, providing a qualitative assessment of each scenario's validity.

The combination of Sentinel-2 and topographic data (S2 and TOPO) provides the most effective scenario for achieving high classification accuracy in a multiclass context. Pizarro et al. (2022) [31] obtained the highest reliability with this same combination of variables using Random Forest. However, their study focused on discriminating against land use and land cover in the Andes, while ours specifically targeted the identification of *Polylepis*.

Discrimination between *Polylepis* and pine was more effective in a 10 m scenario (SC10 and SC12) despite the fact that the metrics Accuracy and the Kappa of scenarios SC4 and SC6 do not have a significant difference. This makes us think that a greater range of radiometric spectra promotes the distinction of *Polylepis*. This finding challenges the assumption that higher spatial resolution automatically translates to better classification results.

The superior performance of the multiclass approach, as opposed to binary, suggests that capturing the variation within *Polylepis* stands, rather than simply the presence or absence, is crucial for accurate mapping. By distinguishing between multiple classes, the model can account for the ecological diversity within *Polylepis* habitats. This approach enables a more detailed understanding of *Polylepis* distribution, which is essential for effective conservation and management. The stands of these forests often exhibit substantial heterogeneity due to environmental gradients, such as altitude, slope, and soil type.

The reduction in variables using RFE is the best scenario (SC10); it was possible to reduce the number of variables from 27 to 11 (8 spectral and 3 topographic) without significantly affecting accuracy and Kappa. The selected variables were MSI, B9, B2, MSR, CCCI, B12, ARI, and NDVI of S2, in conjunction with aspect, slope, and elevation, characterized by the inclusion of elevation and slope as key variables, findings consistent with prior studies [11,23,24,35]. Conversely, scenarios with the lowest reliability (SC1, SC3, and SC7) omitted topographic factors. This underscores the sensitivity of *Polylepis* to altitude (typically above 3000 m) and slope (favoring drainage and sunlight exposure), highlighting the critical role of these variables in its growth.

In classifying *Polylepis* stands, the moisture content of vegetation, indicated by various spectral indices, plays a crucial role in distinguishing these areas. The Moisture Stress Index (MSI), alongside Sentinel-2 bands B9 and B12, is particularly sensitive to water content in the high canopy moisture. Band B2, while mainly used for detecting chlorophyll content, can also provide insight into vegetation health and indirectly reflect moisture levels. Additionally, indices like the Modified Simple Ratio (MSR) and the Anthocyanin Reflectance Index (ARI) highlight specific vegetation properties, with ARI being useful for detecting stress in plants, which can signal varying moisture content. The Canopy Chlorophyll Content Index (CCCI) integrates red-edge bands, useful for capturing subtle changes in chlorophyll and moisture, which estimates chlorophyll content variations within the canopy, offering insights into plant health and vitality. Previous studies, such as Kapari et al. (2024) [64], compared ML algorithms to estimate water stress in maize crops, obtaining that CCCI was an important predictor variable. Finally, the Normalized Difference Vegetation Index (NDVI), while more general in application, supports overall vegetation classification and health monitoring. By combining these indices, a more accurate understanding of *Polylepis* conditions is achievable, aiding in reliable classification by isolating forest areas with specific moisture and health characteristics.

In RFE, we selected the iteration with the highest values of Accuracy and Kappa as the optimal number of variables, even if the differences from other variable sets were not statistically significant. This approach was chosen to maximize the inclusion of variables, allowing for a broader analysis of their relevance in identifying *Polylepis* stands. By prioritizing the highest value, we aimed to capture a comprehensive range of contributing variables, providing a more robust understanding of their relative importance in the classification model. This method enabled a more thorough interpretation of each variable's role, ultimately supporting a refined and potentially more effective identification of *Polylepis* ecosystems.

The ability to detect smaller *Polylepis* stands was constrained by the spatial resolution of S2, which impacts the final result. While Sentinel-2 provides valuable data for large-scale vegetation analysis due to its high spatial resolution and multispectral bands, its 10 to 20 m resolution poses challenges when it comes to identifying and accurately mapping smaller or fragmented *Polylepis* patches. These limitations can lead to an underestimation of the distribution and extent of these smaller stands.

Future endeavors will prioritize ongoing expert data validation to augment the dataset and acquire additional samples for broader coverage of the study area. Furthermore, there will be an exploration of advanced techniques like Neural Networks applied to high spatial resolution imagery such as PS, aiming to enhance classification reliability in remote sensing images.

5. Conclusions

This work demonstrates the effectiveness of combining Sentinel-2 imagery with topographic variables for the detection of *Polylepis* stands. The integration of Sentinel-2 spectral bands and topographic data outperforms the use of higher spatial resolution data from PlanetScope, texture-based variables derived from the NIR spectral band of PlanetScope, or SAR data from Sentinel-1. The importance of spectral and topographic features over spatial resolution in this specific classification task is notably.

Utilizing RFE with smileRandomForest within GEE proved to be effective, reducing the number of variables from 27 to 11 without compromising accuracy. The key variables selected include Sentinel-2 derived indices and bands (NDVI, MSR, MSI, B9, B2, CCCI, B12, and ARI) along with topographic features (elevation, aspect, and slope). The selected variables represent a balanced combination of spectral indices, specific Sentinel-2 bands, and topographic features, indicating that these features are essential for differentiating *Polylepis* stands from other land cover types.

The streamlined feature set includes critical indices sensitive to vegetation health, structure, and moisture content, such as NDVI, MSR, and CCCI, as well as elevation and slope, which are known to influence Polylepis distribution. By focusing on these key variables, the model achieves a high degree of accuracy in detecting Polylepis habitats, which are often found in challenging, high-altitude environments with unique topographic characteristics. The effectiveness of this variable reduction also underscores the value of the

RFE process for refining remote sensing models, enabling a focus on the most informative features while reducing noise from less relevant data.

Moreover, the success of smileRandomForest within GEE highlights the utility of cloud-based platforms for large-scale ecological classification tasks. GEE's extensive data library and computational power facilitate the integration of multispectral, SAR, and topographic datasets, making it possible to replicate and scale this approach for other regions or similar ecological studies. This streamlined approach not only improves the efficiency of data processing and analysis but also makes it feasible to apply the model across large or multi-temporal datasets, supporting long-term monitoring efforts.

Random Forest (RF) has been demonstrated to be a powerful tool for managing imbalanced datasets, especially when distinguishing the *Polylepis* class. Additionally, RF is highly resilient to overfitting, making it particularly effective in complex classification tasks.

The results suggest that the identification of *Polylepis* stands was more effective when combining spectral (S2) and topographic variables, with spectral resolution playing a more critical role than spatial resolution.

Identifying *Polylepis* forests using satellite data is essential as a baseline for applying advanced remote sensing techniques in Land-Use Change monitoring in the Andean region. This approach enables tracking changes over time with diverse data sources, including radar and spectral imagery, to improve classification accuracy and support the establishment of remote monitoring mechanisms. Accurately mapping these unique, high-altitude ecosystems enhances the precision of LULUCF inventories and identifies areas where *Polylepis* forests significantly contribute to carbon sequestration. Developing robust detection methodologies and validating the reliability of remote sensing data for these stands can greatly support national and international reporting requirements.

Author Contributions: Conceptualization, D.P.-P. and L.Á.R.; methodology, D.P.-P.; investigation, D.P.-P. and E.B.-L.; writing—original draft preparation, D.P.-P. and E.B.-L.; supervision, L.Á.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by UNIVERSIDAD DEL AZUAY in the context of investigation of projects 2020-0125 and 2022-0159 denominated "Caracterización de unidades forestales a partir de datos espectrales, espaciales y de relieve a distintas escalas. Aplicación a los bosques andinos del cantón Cuenca (Ecuador)". Fase 2 y 4.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tejedor Garavito, N.; Newton, A.C.; Oldfield, S. Regional Red List Assessment of Tree Species in Upper Montane Forests of the Tropical Andes. *Oryx* **2015**, *49*, 397–409. [CrossRef]
2. Garavito, N.T.; Álvarez, E.; Caro, S.A.; Murakami, A.A.; Blundo, C.; Espinoza, T.E.B.; Torre, M.A. La Evaluación Del Estado de Conservación de Los Bosques Montanos En Los Andes Tropicales. *Ecosistemas* **2012**, *21*, 148–166.
3. Castillo, D.D.; Carrasco, J.C.; Quevedo, L.A.; Ricaurte, C.B.; Gavilanes, A.V.; Borz, S.A. Diversity, Composition and Structure of Andean High Forest in Ecuador, South America. *Bull. Transilv. Univ. Braşov. Ser. II Forest. Wood Ind. Agric. Food Eng.* **2017**, *10*, 1–16.
4. Cuyckens, G.A.E.; Renison, D. Ecología y Conservación de Los Bosques Montanos de Polylepis: Una Introducción al Número Especial. *Ecol. Austral.* **2018**, *28*, 157–162. [CrossRef]
5. Armenteras, D.; Rodríguez, N.; Retana, J.; Morales, M. Understanding Deforestation in Montane and Lowland Forests of the Colombian Andes. *Reg Environ. Chang.* **2011**, *11*, 693–705. [CrossRef]
6. Kleemann, J.; Zamora, C.; Villacis-Chiluisa, A.B.; Cuenca, P.; Koo, H.; Noh, J.K.; Fürst, C.; Thiel, M. Deforestation in Continental Ecuador with a Focus on Protected Areas. *Land* **2022**, *11*, 268. [CrossRef]
7. Balthazar, V.; Vanacker, V.; Molina, A.; Lambin, E.F. Impacts of Forest Cover Change on Ecosystem Services in High Andean Mountains. *Ecol. Indic.* **2015**, *48*, 63–75. [CrossRef]
8. Navarro, G.; Molina, J.A.; Barra, N.D. la Classification of the High-Andean Polylepis Forests in Bolivia. *Plant. Ecol.* **2005**, *176*, 113–130. [CrossRef]

9. Sierra, R.; Calva, O.; Guevara, A. *La Deforestación En El Ecuador, 1990–2018. Factores Promotores y Tendencias Recientes*; Ministerio de Ambiente y Agua del Ecuador: Quito, Ecuador, 2021.
10. Contreras, O. Identificación de La Especie Polylepis Reticulata Mediante Teledetección En Las Zonas Alto Andinas Del Ecuador. Bachelor's Thesis, Escuela Superior Politécnica de Chimborazo, Riobamba, Ecuador, 2019.
11. López, V.L.; Huertas Herrera, A.; Rosas, Y.M.; Cellini, J.M. Optimal Environmental Drivers of High-Mountains Forest: Polylepis Tarapacana Cover Evaluation in Their Southernmost Distribution Range of the Andes. *Trees For. People* **2022**, *9*, 100321. [CrossRef]
12. Caballero-Villalobos, L.; Fajardo-Gutiérrez, F.; Calbi, M.; Silva-Arias, G.A. Climate Change Can Drive a Significant Loss of Suitable Habitat for Polylepis Quadrijuga, a Treeline Species in the Sky Islands of the Northern Andes. *Front. Ecol. Evol.* **2021**, *9*, 661550. [CrossRef]
13. Kivinen, S.; Koivisto, E.; Keski-Saari, S.; Poikolainen, L.; Tanhuanpää, T.; Kuzmin, A.; Viinikka, A.; Heikkinen, R.K.; Pykälä, J.; Virkkala, R.; et al. A Keystone Species, European Aspen (*Populus tremula* L.), in Boreal Forests: Ecological Role, Knowledge Needs and Mapping Using Remote Sensing. *For. Ecol. Manag.* **2020**, *462*, 118008. [CrossRef]
14. Schwager, P.; Berg, C. Remote Sensing Variables Improve Species Distribution Models for Alpine Plant Species. *Basic Appl. Ecol.* **2021**, *54*, 1–13. [CrossRef]
15. Ma, M.; Liu, J.; Liu, M.; Zeng, J.; Li, Y. Tree Species Classification Based on Sentinel-2 Imagery and Random Forest Classifier in the Eastern Regions of the Qilian Mountains. *Forests* **2021**, *12*, 1736. [CrossRef]
16. Xu, D.; Wang, H.; Xu, W.; Luan, Z.; Xu, X. LiDAR Applications to Estimate Forest Biomass at Individual Tree Scale: Opportunities, Challenges and Future Perspectives. *Forests* **2021**, *12*, 550. [CrossRef]
17. Bjerreskov, K.S.; Nord-Larsen, T.; Fensholt, R. Classification of Nemoral Forests with Fusion of Multi-Temporal Sentinel-1 and 2 Data. *Remote Sens.* **2021**, *13*, 950. [CrossRef]
18. Georgieva, M.; Georgiev, G.; Ivanov, M.; Tyufekchiev, K.; Belilov, S.; Dodev, Y.; Glogov, P.; Georgieva, M.L.; Popov, G.; Madzhov, S. Integrated Approach for Health Status Assessment of Urban Forests in the Municipalities of Etropole and Pravets, Bulgaria. *Silva Balc.* **2022**, *23*, 89–102. [CrossRef]
19. Chraibi, E.; Arnold, H.; Luque, S.; Deacon, A.; Magurran, A.E.; Féret, J.B. A Remote Sensing Approach to Understanding Patterns of Secondary Succession in Tropical Forest. *Remote Sens.* **2021**, *13*, 2148. [CrossRef]
20. Luber, A.; Ramachandran, V.; Jaafar, W.S.W.M.; Bajaj, S.; De-Miguel, S.; Cardil, A.; Doaemo, W.; Mohan, M. UAVs for Monitoring Responses of Regenerating Forests under Increasing Climate Change-Driven Droughts—A Review. *IOP Conf. Ser. Earth Environ. Sci.* **2022**, *1167*, 012030. [CrossRef]
21. Vali, A.; Comai, S.; Matteucci, M. Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. *Remote Sens.* **2020**, *12*, 2495. [CrossRef]
22. Vega Isuhuaylas, L.A.; Hirata, Y.; Santos, L.C.V.; Torobeo, N.S. Natural Forest Mapping in the Andes (Peru): A Comparison of the Performance of Machine-Learning Algorithms. *Remote Sens.* **2018**, *10*, 782. [CrossRef]
23. Dostálová, A.; Hollaus, M.; Milenković, M.; Wagner, W. Forest Area Derivation from Sentinel-1 Data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *III–7*, 227–233. [CrossRef]
24. Mngadi, M.; Odindi, J.; Peerbhay, K.; Mutanga, O. Examining the Effectiveness of Sentinel-1 and 2 Imagery for Commercial Forest Species Mapping. *Geocarto Int.* **2021**, *36*, 1–12. [CrossRef]
25. Ma, W.; Yang, H.; Wu, Y.; Xiong, Y.; Hu, T.; Jiao, L.; Hou, B. Change Detection Based on Multi-Grained Cascade Forest and Multi-Scale Fusion for SAR Images. *Remote Sens.* **2019**, *11*, 142. [CrossRef]
26. Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.; Murayama, Y.; Ranagalage, M. Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sens.* **2020**, *12*, 2291. [CrossRef]
27. Xi, Y.; Ren, C.; Tian, Q.; Ren, Y.; Dong, X.; Zhang, Z. Exploitation of Time Series Sentinel-2 Data and Different Machine Learning Algorithms for Detailed Tree Species Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7589–7603. [CrossRef]
28. Akbari, V.; Solberg, S.; Puliti, S. Multitemporal Sentinel-1 and Sentinel-2 Images for Characterization and Discrimination of Young Forest Stands under Regeneration in Norway. *J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 5049–5063. [CrossRef]
29. Liu, Y.; Gong, W.; Hu, X.; Gong, J. Forest Type Identification with Random Forest Using Sentinel-1A, Sentinel-2A, Multi-Temporal Landsat-8 and DEM Data. *Remote Sens.* **2018**, *10*, 946. [CrossRef]
30. Svoboda, J.; Štych, P.; Laštovička, J.; Paluba, D.; Kobliuk, N. Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data—A Case Study of Czechia. *Remote Sens.* **2022**, *14*, 1189. [CrossRef]
31. Pizarro, S.E.; Pricope, N.G.; Vargas-Machuca, D.; Huanca, O.; Ñaupari, J. Mapping Land Cover Types for Highland Andean Ecosystems in Peru Using Google Earth Engine. *Remote Sens.* **2022**, *14*, 1562. [CrossRef]
32. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
33. Vizzari, M. PlanetScope, Sentinel-2, and Sentinel-1 Data Integration for Object-Based Land Cover Classification in Google Earth Engine. *Remote Sens.* **2022**, *14*, 2628. [CrossRef]
34. Rojas, R.; Flexas, J.; Coopman, R.E. Particularities of the Highest Elevation Treeline in the World: Polylepis Tarapacana Phil. as a Model to Study Ecophysiological Adaptations to Extreme Environments. *Flora Morphol. Distrib. Funct. Ecol. Plants* **2022**, *292*, 152076. [CrossRef]
35. Hościło, A.; Lewandowska, A. Mapping Forest Type and Tree Species on a Regional Scale Using Multi-Temporal Sentinel-2 Data. *Remote Sens.* **2019**, *11*, 929. [CrossRef]

36. Tassi, A.; Vizzari, M. Object-Oriented LULC Classification in Google Earth Engine Combining SNIC, GLCM, and Machine Learning Algorithms. *Remote Sens.* **2020**, *12*, 3776. [CrossRef]
37. Ai, J.; Zhang, C.; Chen, L.; Li, D. Mapping Annual Land Use and Land Cover Changes in the Yangtze Estuary Region Using an Object-Based Classification Framework and Landsat Time Series Data. *Sustainability* **2020**, *12*, 659. [CrossRef]
38. Prado, D.P.; Ruiz, L.Á. Comparative of Machine Learning Algorithms and Datasets to Classify Natural Coverage in the Cajas National Park (Ecuador) Based on GEOBIA Approach. *Proc. West Mark Ed. Assoc. Conf.* **2019**, *19*, 9020. [CrossRef]
39. Haq, M.A.; Rahaman, G.; Baral, P.; Ghosh, A. Deep Learning Based Supervised Image Classification Using UAV Images for Forest Areas Classification. *J. Indian Soc. Remote Sens.* **2021**, *49*, 601–606. [CrossRef]
40. Ruiz, L.Á.; Recio, J.A.; Crespo-Peremarch, P.; Sapena, M. An Object-Based Approach for Mapping Forest Structural Types Based on Low-Density LiDAR and Multispectral Imagery. *Geocarto Int.* **2016**, *33*, 443–457. [CrossRef]
41. Bolyn, C.; Michez, A.; Gaucher, P.; Lejeune, P.; Bonnet, S. Forest Mapping and Species Composition Using Supervised per Pixel Classification of Sentinel-2 Imagery. *Biotechnol. Agron. Soc. Environ.* **2018**, *22*, 172–187. [CrossRef]
42. Ligono, L.K.; Okolie, C.J. Integrated Analysis of Mangrove Changes Using the Mangrove Vegetation Index and Random Forest Classification in the Gambia. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *XLVI-M-2–2022*, 153–157. [CrossRef]
43. De Luca, G.; Silva, J.M.N.; Cerasoli, S.; Araújo, J.; Campos, J.; Di Fazio, S.; Modica, G. Object-Based Land Cover Classification of Cork Oak Woodlands Using UAV Imagery and Orfeo Toolbox. *Remote Sens.* **2019**, *11*, 1238. [CrossRef]
44. Georganos, S.; Grippa, T.; Vanhuysse, S.; Lennert, M.; Shimoni, M.; Wolff, E. Very High Resolution Object-Based Land Use-Land Cover Urban Classification Using Extreme Gradient Boosting. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 607–611. [CrossRef]
45. Balha, A.; Mallick, J.; Pandey, S.; Gupta, S.; Singh, C.K. A Comparative Analysis of Different Pixel and Object-Based Classification Algorithms Using Multi-Source High Spatial Resolution Satellite Data for LULC Mapping. *Earth Sci. Inform.* **2021**, *14*, 2231–2247. [CrossRef]
46. Castillo, D.D.; Jara, C.A.; Ricaurte, C.B.; Vaca, B.E.; Quevedo, L.A. Photosynthetic Activity and Canopy Height Model Determined by Uav Rgb and IR Close-Range Remote-Sensing in the High Andean Polylepis Relict Forest, Ecuador. *Bull. Transilv. Univ. Bras. Ser. II For. Wood Ind. Agric. Food Eng.* **2018**, *11*, 1–16.
47. Li, X.; Zheng, Z.; Xu, C.; Zhao, P.; Chen, J.; Wu, J.; Zhao, X.; Mu, X.; Zhao, D.; Zeng, Y. Individual Tree-Based Forest Species Diversity Estimation by Classification and Clustering Methods Using UAV Data. *Front. Ecol. Evol.* **2023**, *11*, 1139458. [CrossRef]
48. Zhou, J.; Chen, X.; Li, S.; Dong, R.; Wang, X.; Zhang, C.; Zhang, L. Multispecies Individual Tree Crown Extraction and Classification Based on BlendMask and High-Resolution UAV Images. *J. Appl. Remote Sens.* **2023**, *17*, 016503. [CrossRef]
49. Minga, D.; Ansaloni, R.; Verdugo, A.; Ulloa, C.U. Flora Del Páramo Del Cajas, Ecuador. Bachelor's Thesis, Universidad del Azuay, Cuenca, Ecuador, 2016.
50. Ulloa, C.; Álvarez, S.; Jørgensen, P.; Minga, D. Guía de 100 Plantas Silvestres Del Páramo Del Parque Nacional Cajas. 2008. Available online: https://www.mobot.org/MOBOT/research/paramo/pdf/GuiaPlantas.pdf (accessed on 7 October 2024).
51. FAO; UNEP. *El Estado de Los Bosques Del Mundo 2020*; FAO and UNEP: Rome, Italy, 2020; ISBN 978-92-5-132421-9. [CrossRef]
52. Tenesaca, C.; Quindi, T.; Delgado, G.; Toledo, E.; Delgado, O. Generación Del Mapa de Cobertura y Uso Del Suelo de La Provincia Del Azuay. *Inst. De Estud. De Régimen Secc. Del Ecuad. IERSE* **2017**, *73*, 23–37. [CrossRef]
53. SIGTIERRAS Ortofotografía Del Ecuador. El País Desde Otra Perspectiva. *Minist. De Agric. Y Ganad.* **2017**, *1*, 88.
54. Kumar, L.; Mutanga, O. Google Earth Engine Applications since Inception: Usage, Trends, and Potential. *Remote Sens.* **2018**, *10*, 1509. [CrossRef]
55. Sentinel-2 MSI: MultiSpectral Instrument, Level-2A. Available online: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR#description (accessed on 30 March 2021).
56. Nandy, S.; Srinet, R.; Padalia, H. Mapping Forest Height and Aboveground Biomass by Integrating ICESat-2, Sentinel-1 and Sentinel-2 Data Using Random Forest Algorithm in Northwest Himalayan Foothills of India. *Geophys. Res. Lett.* **2021**, *48*, e2021GL093799. [CrossRef]
57. Arekhi, M.; Goksel, C.; Sanli, F.B.; Senel, G. Comparative Evaluation of the Spectral and Spatial Consistency of Sentinel-2 and Landsat-8 OLI Data for Igneada Longos Forest. *ISPRS Int. J. Geoinf.* **2019**, *8*, 56. [CrossRef]
58. IDB—Index DataBase. Available online: https://www.indexdatabase.de/ (accessed on 7 May 2024).
59. Praticò, S.; Solano, F.; Di Fazio, S.; Modica, G. Machine Learning Classification of Mediterranean Forest Habitats in Google Earth Engine Based on Seasonal Sentinel-2 Time-Series and Input Image Composition Optimisation. *Remote Sens.* **2021**, *13*, 586. [CrossRef]
60. Mishra, A.P.; Rai, I.D.; Pangtey, D.; Padalia, H. Vegetation Characterization at Community Level Using Sentinel-2 Satellite Data and Random Forest Classifier in Western Himalayan Foothills, Uttarakhand. *J. Indian Soc. Remote Sens.* **2021**, *49*, 759–771. [CrossRef]
61. Grabska, E.; Hostert, P.; Pflugmacher, D.; Ostapowicz, K. Forest Stand Species Mapping Using the Sentinel-2 Time Series. *Remote Sens.* **2019**, *11*, 1197. [CrossRef]
62. Pratama, B. Performing Feature Selection in Google Earth Engine (GEE) Using Recursive Feature Elimination (RFE) | by Pratama Bima | Artificial Intelligence in Plain English. Available online: https://ai.plainenglish.io/performing-feature-selection-in-google-earth-engine-gee-using-recursive-feature-elimination-rfe-0f5ba2789c4c (accessed on 22 June 2024).

63. Gerón, A. *Hands-on Machine Learning with Scikit-Learn, Keras & Tensorflow*, 2nd ed.; O'Reilly Media: Sebastopol, CA, USA, 2019; ISBN 9781492032649.
64. Kapari, M.; Sibanda, M.; Magidi, J.; Mabhaudhi, T.; Nhamo, L.; Mpandeli, S. Comparing Machine Learning Algorithms for Estimating the Maize Crop Water Stress Index (CWSI) Using UAV-Acquired Remotely Sensed Data in Smallholder Croplands. *Drones* **2024**, *8*, 61. [CrossRef]