International Journal of Applied Earth Observation and Geoinformation 152 (2026) 105454

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

# A machine learning framework for predicting agricultural terrace suitability using positive and unlabelled datasets: an application for the Troodos Mountains, Cyprus

Aman Kumar Meena [a], Christos Zoumides [a], Hakan Djuma [a], Ioannis Sofokleous [a], Corrado Camera [b], Adriana Bruggeman [a]

[a] Energy, Environment and Water Research Centre, The Cyprus Institute, Aglantzia, Lefkosia 2121, Cyprus
[b] Dipartimento di Scienze Della Terra "A. Desio", Università Degli Studi di Milano, Via Mangiagalli 34, 20133 Milan, Italy

ARTICLE INFO

*Keywords:*
Mountain terraces
Positive and unlabelled learning
XGBoost
SHAP
Google Earth engine

ABSTRACT

Terraces are a defining feature of Mediterranean mountain landscapes, enabling agriculture on steep slopes while providing multiple ecosystem services. Land suitability analysis (LSA) can guide authorities and land users to sustainably manage and expand these environments. It typically requires fully labelled datasets, but in many real-world applications only a fraction of positive examples is available, with the rest unlabelled. This study aims to present an integrated predictive modelling framework that combines GIS with data-driven Machine Learning (ML) techniques, capable of learning from Positive and Unlabelled (PU) datasets for LSA. The proposed framework was applied to develop a terrace suitability map for Cyprus' Troodos Mountains. A 5-m DEM was processed to extract the mountain area, with elevation $\geq$ 500 m and slopes $\geq$ 15%, defining the study area. Crop plots registered under the Single Area Payment Scheme of the European Common Agricultural Policy were used to classify the study area into Terrace-Present (TP) and Terrace-Absent (TA) cells, with TP serving as labelled positive and TA as unlabelled samples. A two-step ML approach was applied, first identifying reliable negatives from TA cells, then using these with TP cells for suitability prediction. The developed PU-classifier was evaluated under the selected completely-at-random (SCAR) assumption, achieving a Recall of 84.6%, Precision of 81.5% and an F1 score of 83%. Feature importance analysis identified land cover, terrain slope and tree cover density as the most influential parameters for terrace suitability. Comparative analysis between 2017 and 2024 revealed abandonment of terraced agricultural land (29% decrease) as well as revitalisation (12% increase). The resulting suitability map and accompanying data layers are accessible through a Google Earth Engine application, aiming to support informed decision-making for sustainable landscape planning.

## 1. Introduction

Terracing is a long-established conservation practice in mountainous regions, offering benefits such as improved water management, increased arable land, crop diversification, and slope stabilization (Chen et al., 2017; Deng et al., 2021). It also provides a range of ecosystem services, contributing both directly and indirectly towards achieving sustainable development goals, as summarized by Zoumides et al. (2025). Despite their significant value, terraced fields are often underutilized (Chen et al., 2022). Over the past decades, many terraces have been abandoned due to socioeconomic shifts, high labour demands, and maintenance costs, leading to a progressive increase in land degradation and the loss of soil functions in mountain environments (Dax et al., 2021; Tarolli, 2018; Zoumides et al., 2022). However, in recent years, the adoption of mechanised construction has reduced these constraints and contributed to the development of new terraced systems (Ramos et al., 2007; FAO, 2019; Franco et al., 2020; Romeo et al., 2021). As a result, the protection and rehabilitation of terraced agriculture have gained renewed interest (Zoumides et al., 2022), supported by various policy initiatives that have emerged in the early 21st century to conserve these mountain terraced landscapes, as summarized by Lasanta et al. (2013). The Troodos Mountains in Cyprus represent a typical example of such landscapes, where traditional terraces face increasing pressures from land abandonment (Zoumides et al., 2025) and associated soil erosion risks (Djuma et al., 2017; Camera et al., 2018), underscoring the need for evidence-based tools to support sustainable planning and resource allocation. To advance such conservation efforts, Land Suitability Analysis (LSA) can offer scientifically-based guidance to authorities, decision-makers, and land users for the effective implementation of agricultural policies and land use frameworks (Geng et al., 2019; Alonso et al., 2025).

Modern machine learning (ML) algorithms are increasingly

https://doi.org/10.1016/j.jag.2026.105454
Received 31 October 2025; Received in revised form 12 June 2026; Accepted 1 July 2026

employed in LSA and have been reported in some applications to improve predictive performance and reduce subjectivity relative to traditional multi-criteria decision-making approaches (Radočaj & Jurišić, 2022). Models such as Random Forest (RF) (Sahu et al., 2023; Singh et al., 2022), Convolutional Neural Networks (CNN) (Carranza-García et al., 2019; Luan et al., 2023), K-Nearest Neighbors (KNN) (Ganesan et al., 2021), Extreme Gradient Boosting (XGBoost) (Batunacun et al., 2021), and Support Vector Machines (SVM) (Taghizadeh-Mehrjardi et al., 2020) have been widely applied for land classification and suitability prediction. Numerous comparative studies focusing on different supervised machine learning algorithms have indicated that ensemble methods, particularly RF and XGBoost, consistently outperform other algorithms due to their advanced learning strategies of bootstrap aggregating (for RF) and sequential boosting (for XGBoost) (Shawon et al., 2025; Xing et al., 2022; Sinchana et al., 2024; Negussie et al., 2024; Karongo et al., 2025, Lekshmi et al., 2025). However, direct comparisons between RF and XGBoost have produced mixed outcomes. For instance, Xing et al. (2022) compared RF and XGBoost for tea cultivation suitability for an area of 139.03 $km^2$ using 12 covariates, reporting that RF achieved slightly higher performance (AUC = 0.86) than XGBoost (AUC = 0.82). Georgiades et al. (2025) compared XGBoost and RF using a global climate dataset spanning 18 years (290 million data points) with constant hyperparameters. They reported that RF required three times longer training and over three times more memory than XGBoost. Moreover, predictions generated by XGBoost were more closely aligned with the ground truth than those of RF. The study also emphasized the importance of incremental learning for handling datasets of this scale, which is supported by XGBoost but not by RF. Lekshmi et al. (2025) also reported the superiority of XGBoost over RF in handling non-linear patterns, complexity, and class imbalance in the dataset due to its capabilities of incorporating regularization parameters (L1 and L2) and boosting strategy. These findings suggest that model performance is context-specific, depending on dataset characteristics, feature selection, and tuning strategies (Torgbor et al., 2023; Mali et al., 2023).

While supervised ML models have demonstrated strong performance in land suitability classification, they have typically relied on fully labelled datasets with clearly defined positive and negative examples (Saunders and Freitas, 2025). However, in many real-world applications, this condition is not satisfied, and only a set of the positive examples is available while the remaining majority of instances are unlabelled (i.e., unclear if they are positive or negative) (Saunders and Freitas, 2025). For instance, in landslide susceptibility mapping, relevant inventories provide historical records of landslide events and are treated as positive examples. However, the absence of recorded landslide events does not necessarily indicate an absence of susceptibility (Zhao et al., 2023). Therefore, such instances should be considered unlabelled rather than negative. To address such issues, numerous PU (positive and unlabelled) learning techniques, i.e. specialized cases of semi-supervised learning, have been proposed, as summarized by Bekker and Davis (2020). One of the techniques is a two-step method that consist of identifying reliable negative examples from the unlabelled data and utilizing positive labelled and reliable negative examples to classify remaining unlabelled data using traditional supervised classification (Bekker and Davis, 2020; Zheng et al., 2019). Zhang et al. (2024a) developed a landslides susceptibility map by integrating traditional ML algorithms including RF, KNN and Decision Tree (DT) with the PU-learning technique under the Selected Completely at Random (SCAR) assumption, as introduced by Elkan and Noto (2008). They concluded that PU learning substantially improves the Area Under Curve and Recall, leading to a more conservative LSM compared to binary classification methods. Li et al. (2022) compared the Presence and Background Learning with Constraints (PBLC) PU-learning technique to a traditional Artificial Neural Network (ANN) for flood susceptibility prediction. They trained a binary classifier using case-control positive and unlabelled samples without truly labelled negative data, alongside 14 environmental covariates. The study concluded that PBLC provides better calibrated probabilistic predictions, more accurate binary classification, and more reliable urban flood susceptibility mapping than the traditional ANN. Numerous other PU learning applications are available in different fields such as in mapping groundwater potential (Bi et al., 2025), classification of remote sensing data (Desloires et al., 2022), ecological studies (Li et al., 2011), bioinformatics (Yang et al., 2012) and computer vision (Wei et al., 2020). Although PU-learning approaches are widely used, they still face model validation challenges, as the absence of negative examples complicates the assessment of classification performance. The SCAR assumption provides a way to overcome this challenge by assuming that the labelled positive samples are randomly drawn from the entire positive set, allowing reliable evaluation of classifier performance despite the absence of true negative data (Saunders and Freitas, 2025).

Despite the methodological developments, no mapping applications of PU-learning for mountain terrace suitability have been reported in the scientific literature. Similar to other PU-learning problems, the absence of terraced land does not necessarily imply unsuitability of land for terracing and should be treated as unlabelled rather than negative. Therefore, this study aims (i) to present an integrated predictive modelling framework that combines GIS with data-driven ML techniques, capable of learning from positive and unlabelled datasets for terrace suitability analysis, (ii) to generate a high-resolution terrace suitability map for the Troodos Mountains, Cyprus (iii) to analyse the environmental, regulatory and socio-economic parameters that affect the suitability of land for mountain terraces and (iv) to analyse recent land use dynamics (i.e., abandonment and expansion of mountain terraces).

## 2. Methodology

This study was structured in six sequential steps (Fig. 1). First, elevation and slope criteria were used to delineate the mountain area. Second, within this area, all crop plots were assumed to be terraced land. Following this assumption, areas occupied by these crop plots were classified as Terrace-Present (TP) and areas without crop plots were classified as Terrace-Absent (TA). Third, 16 covariates were identified and a correlation analysis was performed to reduce redundancy among variables. Fourth, a hybrid framework of PU-learning and supervised XGBoost classification was adopted to predict the suitability of terrace agriculture using the refined set of non-redundant covariates. The resulting continuous suitability scores were then classified into four distinct suitability classes using the *k*-means clustering method. Fifth, the SHapley Additive exPlanations (SHAP) analysis was performed to evaluate the influence of individual covariates on land suitability. Finally, agricultural crop plots of 2017 and 2024 were utilized to determine the abandonment and expansion of these environments and corresponding suitability classes. The following sections provide a detailed description of each step.

### 2.1. Study area

Cyprus is the third largest island in the Mediterranean with an area of 9,251 $km^2$ and lies between 34°–36°N and 32°–35°E. The Troodos Mountains dominate the centre of Cyprus and cover 40% of the island area (Fig. 2). It comprises an ophiolite complex surrounded by sedimentary formations (Christofi et al., 2020). It has a mean slope gradient of 31% with its highest summit at 1,951 m a.s.l. (Christofi et al., 2020). The island exhibits significant variation in average annual rainfall along its elevation gradient, ranging from 500 mm in the lower foothills to 1,100 mm near the peak of the Troodos Mountains (Camera et al., 2014). Agricultural terraces in the Troodos Mountains have traditionally been constructed by cutting and filling of steep slopes (Zoumides et al., 2017). The Cyprus Rural Development Program 2014–2020 (Ministry of Agriculture, Natural Resources and Environment, 2021) classifies the

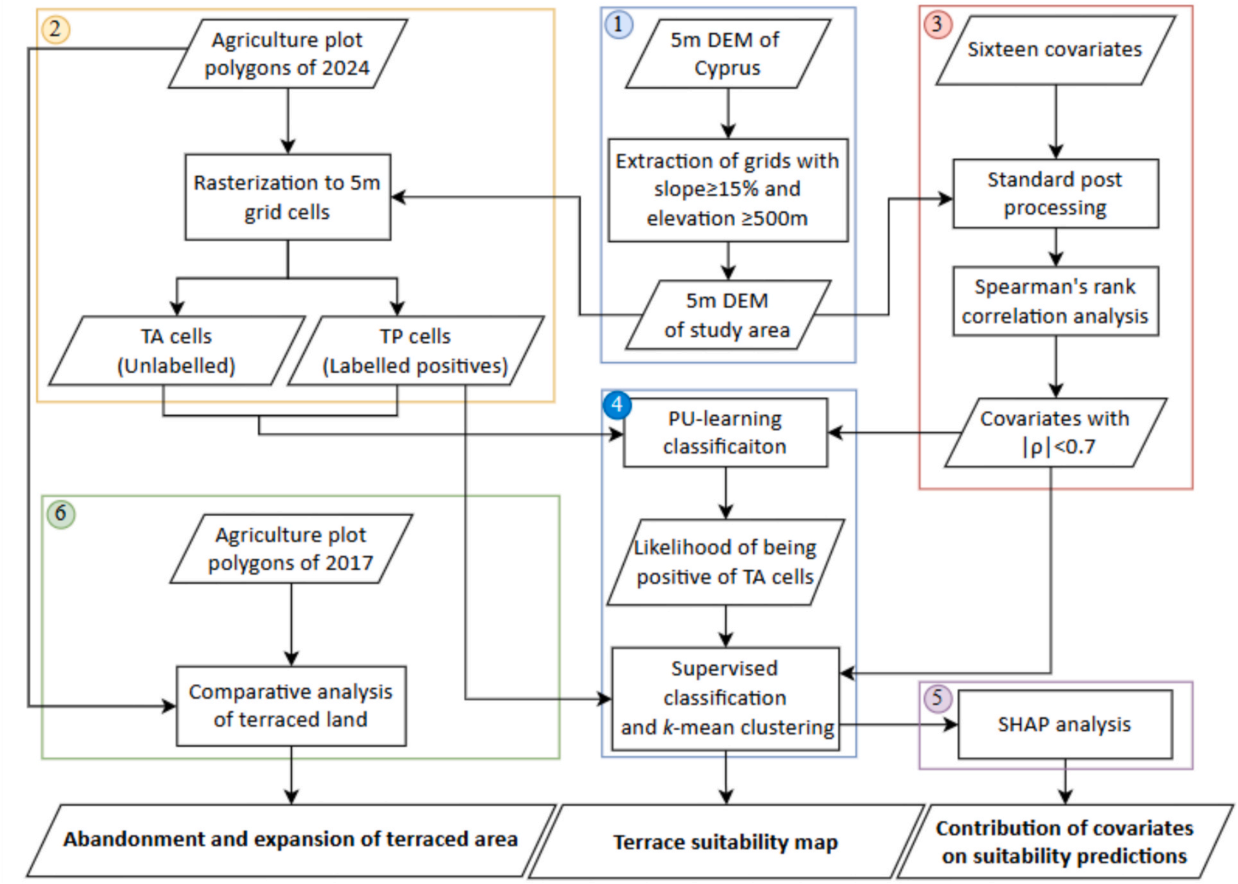


**Fig. 1.** Conceptual workflow summarizing the six methodological steps adopted for suitability mapping, feature importance analysis and recent land use dynamics; rectangles denote computational processes and parallelogram show data inputs/outputs.

areas with altitude higher than 800 m, or between 500 and 800 m with slopes $\geq$ 15% as mountain areas. This was defined in line with EU regulation 1305/2013, article 32 and Annex III (European Parliament and Council of the European Union, 2013), which characterizes mountain areas by the existence of a short growing season (< 180 days) or steep slopes ($\geq$15%) or a combination of both. Numerous researchers have reported varying values for the slopes of the terraced land (Stanchi et al., 2012). According to the Food and Agriculture Organization (FAO, 1988), manually constructed bench terraces are recommended for slopes ranging from 7 to 25 degrees (12%–47%), while machine-constructed terraces are recommended for slopes between 7 and 20 degrees (12%–36%). To align with both Cyprus Rural Development Program and FAO guidelines, this study delineates areas on Troodos Mountains having elevation $\geq$ 500 m and slopes $\geq$ 15%, assuming that these topographic thresholds represent the minimum conditions for the feasibility of bench terrace construction in mountain areas (Step 1).

### 2.2. Raster classification of study area

Crop-plot polygons, i.e., agricultural areas registered in 2024, under the Single Area Payment Scheme from farmers' applications for direct subsidies, administered by the Cyprus Agricultural Payment Organization (CAPO), were used to map TP land (Step 2). These polygons were rasterized, at both 5-m and 30-m resolution, with each grid cell assigned the crop type corresponding to the majority area of the underlying polygon. A comparative analysis of the 5-m and 30-m DEMs was conducted to evaluate the effectiveness of DEM resolution for LSA. The analysis was conducted for the area above 500 m elevation and classified the agricultural land into two slope categories, $\geq$15% and < 15%. The DEM resolution that produced a total rasterized agricultural area closest to the original polygon area was selected. Based on this assessment, the 5-m DEM was used for the LSA (Section 3.1).

### 2.3. Description of covariates and correlation analysis

Sixteen covariates were derived using environmental, regulatory and socio-economic datasets, as summarized in Table 1. The 5-m DEM was utilized to extract topographic parameters, including elevation, slope, aspect, curvature and upstream contributing area. Soil pH, soil organic carbon (SOC), and sand and clay content, were obtained from SoilGrids (Poggio et al., 2021) at three depth intervals (0–5 cm, 5–15 cm, and 15–30 cm). These data were used to compute a depth-weighted average representing the top 0–30 cm of the soil profile. The weighted averages for sand and clay content were subsequently classified into texture classes using the HYPRES soil texture classification system (Moeys, 2024). Population density was derived for each community area, expressed as inhabitants per km$^2$, rather than being computed for individual grid cells. Additionally, Euclidean distance from each grid cell to the nearest village centre was calculated to produce a spatial layer representing proximity to settlement areas. The other covariates did not require any significant pre-processing.

A standardized post-processing workflow was applied to all the covariates, ensuring proper alignment and avoiding pixel shift. This workflow includes: (1) transforming the derived covariate layers into raster grids based on the majority area (applicable only to datasets originally in shapefile format), (2) defining a uniform projection system (WGS 1984 UTM Zone 36 N), (3) resampling to a 5-m resolution using the nearest neighbour method, and (4) using the 5-m DEM of the study area as a reference for both snap raster and masking layer. This resampling was used solely to align all covariates with the 5-m DEM grid and

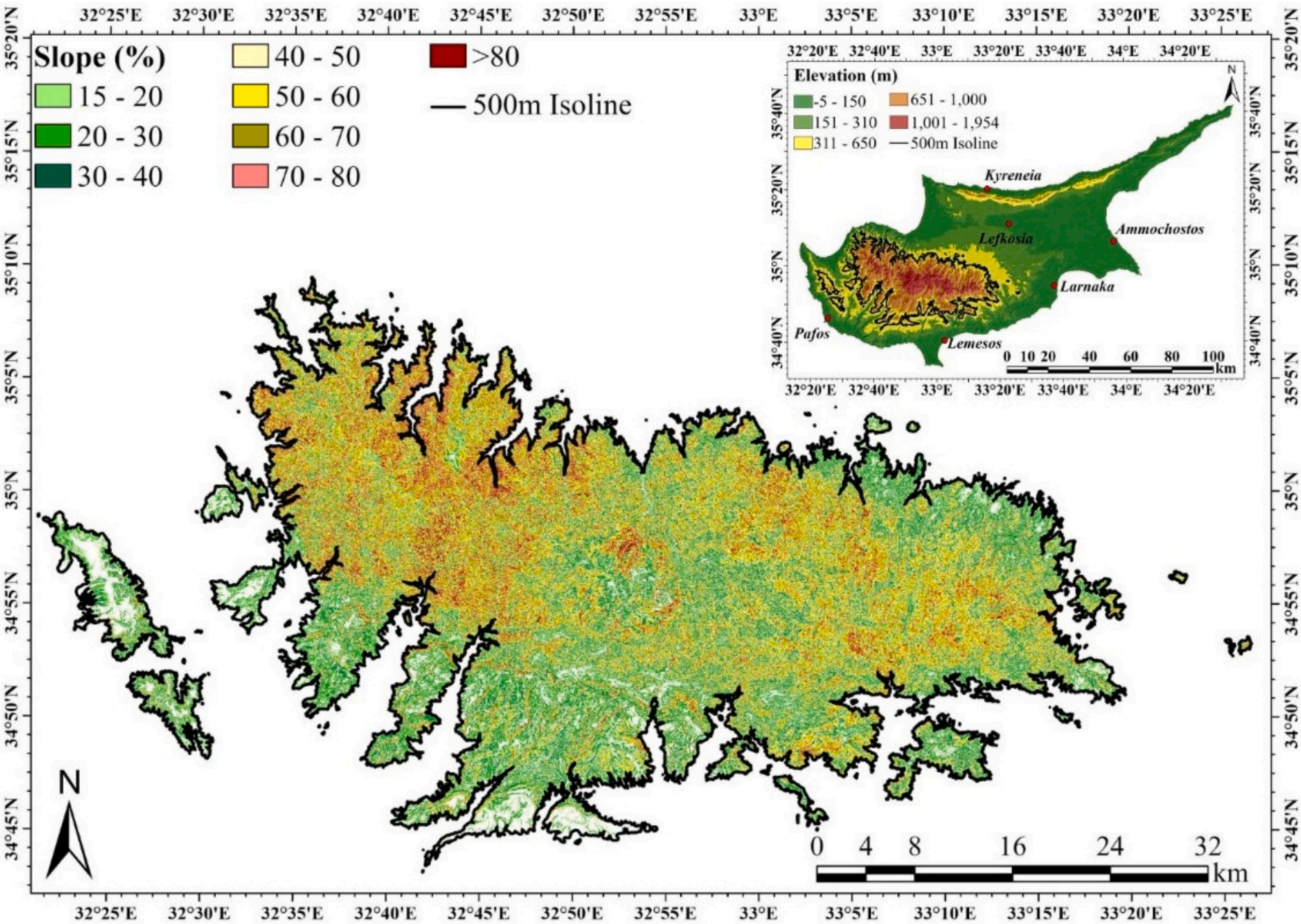


**Fig. 2.** Elevation map of Cyprus with the 500 m isoline and slope map of areas above 500 m, highlighting the Troodos Mountains, delineated using the 5-m DEM (Kyriakou, 2022).

crop-plot raster, without introducing any additional spatial detail beyond the native resolution of the coarser input datasets.

To evaluate pairwise relationships among the identified continuous covariates, Spearman rank correlation analysis was performed. The resulting correlation (ρ) matrix illustrates the direction and strength of monotonic relationships between variables. Similar to Maharjan et al. (2021), covariates with $|\rho| > 0.7$ were excluded from the ML simulations to avoid multicollinearity, reduce redundancy and avoid the risk of overfitting (Step 3).

## *2.4. Model training and prediction workflow*

The dataset contained approximately 64.4 million records covering the whole study area at a resolution of 5 m. Each record contains covariates (listed in Table 1, excluding highly correlated variables) and is assigned to either TP or TA. Among these, approximately 62.1 million records correspond to TA, while around 2.3 million records represent TP. Given that the presence and associated attributes of existing terraces (2.3 million records) are known, all TP records can be used as positive examples for terrace suitability (herein referred to as positive). However, the remaining 62.1 million TA records lack confirmed labels (herein referred to as unlabelled), as they may include both suitable (positive examples) and unsuitable (negative examples) locations for terracing. Considering the absence of confirmed negative examples, a two-step ML framework was adopted to model terrace suitability.

### *2.4.1. Pu-learning classification of unlabelled datasets*

The PU learning method proposed by Elkan and Noto (2008) offers a semi-supervised approach to estimate the likelihood that unlabelled data belong to the positive class (Step 4). This method relies on the Selected Completely At Random (SCAR) assumption, which states that the positive examples in the labelled set are representative of all positive examples (both labelled and unlabelled). In other words, the labelled positive examples form a randomly selected subset of all positive examples (Saunders and Freitas, 2022). Let $x \in R^m$ denote the feature vector of an example, described by $m$ number of covariates, representing a 5-meter grid cell of the study area. Let $y \in \{0, 1\}$ be the binary class for each example, where $y = 1$ indicates a positive example (suitable for terracing) and $y = 0$ indicates a negative example (not suitable for terracing). Additionally, let $s \in \{0, 1\}$ be a binary indicator with $s = 1$ indicating a labelled example and $s = 0$ indicating an unlabelled example. Under the SCAR assumption, Elkan and Noto (2008) proved a lemma to estimate a traditional classifier $f(x) = \Pr(y = 1|x)$, which models the probability of an example belonging to the positive class. This is done through a non-traditional classifier $g(x) = \Pr(s = 1|x)$, which estimates the probability of an example being labelled, and a constant $c = \Pr(s = 1|y = 1)$ indicating the probability that a positive example is labelled. The relationship between these quantities is given by (Elkan and Noto, 2008):

$$f(x) = \frac{g(x)}{c} \tag{1}$$

To address the absence of confirmed negative examples, the Elkan

**Table 1**
Environmental, regulatory and socio-economic covariates for suitability analysis of terraced land.

| Derived covariates | Unit | Type* | Dataset | Spatial resolution | Year | Source |
|---|---|---|---|---|---|---|
| Elevation | m | Continuous | Digital elevation model | 5 m and 30 *m* | 2014 and 2019 | Kyriakou, 2022 and ASTER Global Digital Elevation Map https.//asterweb.jpl.nasa.gov/gdem.asp |
| Slope | % | Continuous | | | | |
| Aspect | ° | Continuous | | | | |
| Upstream area | $m^2$ | Continuous | | | | |
| Curvature | 1/100 *m* | Continuous | | | | |
| Land cover | – | Categorical (21 classes) | CORINE land cover | 100 *m* | 2018 | Copernicus land monitoring services https.//land.copernicus.eu/en |
| Planning zones | – | Categorical (82 classes) | Land zonation | 8,171 polygons | 2017 | Ministry of Interior https://www.gov.cy/moi/en/ |
| Tree cover density | % | Continuous | Copernicus canopy area | 10 *m* | 2018 | Copernicus Land Monitoring Services https.//land.copernicus.eu/en |
| Geological formations | – | Categorical (31 classes) | Geological map of Cyprus | 1.250,000 | 1995 | Geological Survey Department https.//https://www.moa.gov.cy/ |
| Depth-weighted average for soil pH (0–30 cm) | – | Continuous | SoilGrids (soil pH, soil organic carbon, % of sand and clay) | 250 *m* | 2022 | ISRIC – world soil information https.//https://www.isric.org/ |
| Depth-weighted average for soil organic carbon (0–30 cm) | % | Continuous | | | | |
| Depth-weighted average for soil texture (0–30 cm) | – | Categorical (3 classes) | | | | |
| Average annual precipitation | mm | Continuous | CYOBS (precipitation and temperature indicators) | 1 km | 1980 to 2019 | Camera et al., 2014 and Sofokleous et al., 2021 |
| Average annual temperature | °c | Continuous | | | | |
| Population density | count/ $km^2$ | Continuous | Community boundaries | 615 polygons | 2018 | Department of Lands and Surveys https.//portal.dls.moi.gov.cy/en |
| | | | Population census | 615 points | 2021 | Statistical Service of Cyprus https.//https://www.cystat.gov.cy |
| Distance of terraces from centre of village | m | Continuous | Village centre | 615 points | 2010 | Department of Town planning and Housing http.//https://www.moi.gov.cy/ |

* For categorical datasets, only the classes present in the study area are reported (e.g., 21 out of 44 CORINE land cover classes).

and Noto (2008) PU learning method was implemented in Python, by wrapping an XGBoost classifier within the ElkanotoPuClassifier. The dataset was randomly divided into training and validation sets, with 80% of the samples used to train the model and 20% reserved for independent validation to monitor performance. Hyperparameter tuning was performed by testing multiple combinations of model parameters and evaluating their performance on the validation dataset, with the configuration yielding the highest validation accuracy selected for the final model. Accordingly, the underlying XGBoost model was configured to fit 500 trees with a maximum depth of 15 nodes per tree. A learning rate of 0.1 was selected to promote gradual and stable convergence. L2 regularization was applied to reduce overfitting, with L1 regularization set to zero. Furthermore, a gamma value of 1 was used to enforce a minimum loss reduction for further splitting of leaf nodes, improving split quality (Tarwidi et al., 2023;). Similar to Pawłuszek-Filipiak and Lewandowski (2025), the class imbalance between the positive and unlabelled data was addressed using the scale_pos_weight. This value was set to 4. Additionally, 20% of the labelled positive examples was used as a held-out set in the ElkanotoPuClassifier for the estimation of label frequency ″c" (Bekker and Davis, 2020). Training was conducted on a single node (cpus-per-task = 64) of a High-Performance Computing Facility (HPCF), equipped with dual Intel(R) Xeon(R) Gold 6248 processor and 94 GB of RAM, utilizing a total of 1,172 CPU hours.

Traditional supervised classifiers are typically evaluated using metrics such as Precision, Recall, and F1-score. However, in PU-learning, the absence of true class labels makes it infeasible to compute these metrics directly. Nevertheless, under the SCAR assumption, the classifier's behaviour on the labelled positive set is representative of its behaviour on the entire positive set (Saunders and Freitas, 2025). As theorized by Tabatabaei et al. (2021), let, $M := \{1, \ldots, n\}$ be the set of all examples, $P \subseteq M$ be the set of all positive examples, $S \subseteq P$ be the set of all labelled positive examples and $U \subseteq M \backslash S$ be the set of unlabelled examples which may contain both positive and negative examples. Consider a classifier trained using $S$ and U. According to the SCAR assumption, the expected fraction of predicted positives in the labelled positive set is to be the same as the fraction of predicted positives in total positive examples. Therefore, an estimation of $P$ can be obtained which can further be utilized to calculate $\rho'$ (i.e., $\rho' = S/P$), as $S$ is already known. This estimate $\rho'$ can be used to modify the standard evaluation metrics for PU-learning. These equations are presented in Table 2.

*TP*, *FN*, and *FP* represent the true positive, false negative and false positive predictions; $P_1$ is the number of positive examples that are also predicted positive (equivalent to *TP*); *P* is the total positive examples; $M_1$ is the total number of predicted positive examples; $S_1$ is the total predicted positive from the labelled positive set; $S$ is the labelled positive examples.

The trained model was evaluated on the 20% validation set by computing performance metrics (Table 2) under the SCAR assumption. This PU-learning classification resulted in a continuous probability distribution for the unlabelled dataset indicating the likelihood that an example belongs to the positive class. This was further used to extract a set of reliable negatives as described in the following section.

### 2.4.2. Supervised learning with XGBoost

To obtain a sample of reliable negative examples from the unlabelled data, all grid cells with predicted probabilities (of being positive) < 0.3 were stratified into three probability bins: 0.0–0.1, 0.1–0.2, and 0.2–0.3. Based on these bins, three datasets were constructed, while preserving the original class imbalance of each bin. The first dataset included only grid cells from the 0.0–0.1 bin; the second combined grid cells from the 0.0–0.1 and 0.1–0.2 bins; and the third included grid cells within the 0.0–0.3 range, incorporating all three bins. Each dataset was then

**Table 2**
Evaluation metrics for PU-learning under SCAR (Tabatabaei et al., 2021).

| Metrics | Supervised learning | Modified under SCAR |
|---|---|---|
| Recall | $\frac{TP}{TP+FN} = \frac{P_1}{P}$ | $\frac{S_1}{S}$ |
| Precision | $\frac{TP}{TP+FP} = \frac{P_1}{M_1}$ | $\frac{S_1}{\rho'.M_1}$ |
| F-1 score | $2\left[\frac{Recall \times Precision}{Recall + Precision}\right]$ | $2\left[\frac{\rho'.P_1}{\rho'.M_1 + S}\right]$ |

combined with the positive examples to form three complete sets, referred to as Set 1, Set 2, and Set 3. For each set, 80% of the data was used for training, and the remaining 20% was reserved for evaluating model performance. Separate XGBoost classifiers were trained on each set using identical hyperparameters and sample sizes, ensuring that the only variable affecting model performance was the proportion of negative examples included from the probability bins. This approach minimized the potential confounding factors and allowed a fair assessment for the impact of bin selection on classification performance. In addition to the standard random-split evaluation, a cluster-based spatial cross-validation (CV) approach was implemented to assess model robustness across space. Following the recommendations of Sun et al. (2023), the study area was partitioned into 10 spatial clusters using *k*-means clustering based on the x and y coordinates of the grid cells. For each set, five repetitions of the spatial-CV were performed. In each repetition, seven clusters were used for training, while the remaining three clusters were used for validation.

Following the evaluation of these sets using the Precision, Recall, F1-score and Area Under the Curve (AUC) under both random-split and spatial-CV, the bin combination that yielded the best performance was selected as the reliable negative examples. The classifier trained on this selected set was then used to predict suitability probabilities for each grid cell across the study area. To enable practical interpretation and support spatial planning, the resulting continuous probabilities were further classified into discrete land suitability classes. The optimal number of classes was determined using *k*-means clustering applied to the predicted probability values. Cluster solutions ranging from 2 to 10 were evaluated by computing the within-cluster sum of squared errors (SSE), also referred to as inertia (Yuan and Yang, 2019). The Elbow Method was applied to identify the point at which adding additional clusters resulted in minimal reduction in SSE. The number of clusters corresponding to this point was then selected for the final classification (Step 4).

### 2.5. Feature importance analysis

To evaluate the relative contribution of input features (covariates) to the suitability predictions, a feature importance analysis was conducted using SHapley Additive exPlanations (SHAP) technique (Step 5). This is based on cooperative game theory and quantifies the contribution of each feature to the model's output through SHAP values. The TreeExplainer algorithm was employed to compute SHAP values for the XGBoost model. The SHAP values explain how each feature additively shifts the model's log-odds prediction from the base value.

To elaborate, let, $x = (x_1, x_2, x_3, \cdots, x_m)$ represent an individual grid of the dataset defined by its $m$ features. The XGBoost model made a prediction $f'(x)$ in log-odds space, representing the suitability of terracing for the given feature. This prediction can be additively expressed as (Lundberg and Lee, 2017):

$$f'(x) = \varnothing_{base} + \sum_{i=1}^{m} \varnothing_{x_i} \tag{2}$$

where $\varnothing_{base}$ is the base value, defined as the expected model prediction in log-odds, when no feature information is available. Mathematically, it can be also defined as mean prediction of the model when all features are marginalized out. $\varnothing_{x_i}$ is the SHAP value for the feature $i$ quantifying its marginal contribution to the prediction $f'(x)$ relative to $\varnothing_{base}$.

SHAP values enable both global and local interpretation of feature contributions to model predictions. Global interpretation reveals how features influence model predictions across the entire dataset, while local interpretation explains the contribution of each feature to individual predictions. Global feature importance was assessed using the mean absolute SHAP value plot, ranking features by their average impact on model predictions. Additionally, a SHAP beeswarm plot was generated to visualize both the magnitude and direction of feature contributions, indicating how high or low feature values drive predictions toward positive or negative outcomes. SHAP dependence plots were also illustrated for the top three most influential features, highlighting their individual effects and interactions, capturing both local contributions and global feature trends in the suitability analysis.

The land suitability map was intersected with all covariates in ArcGIS using raster algebra. The intersected grids were post-processed by grouping categorical covariates into classes and continuous covariates into ranges. For each covariate, the proportion of suitability classes within these groups was quantified and analysed to assess their distribution with respect to terrace suitability. To assess the influence of dominating variable, a second analysis was also performed after removing the top contributing covariate, replicating the similar ML framework.

### 2.6. Comparative analysis of terraced land for recent terrace dynamics

To assess recent changes in terraced agriculture, crop plot data from 2017 and 2024 were utilized (Step 6). A consistent post processing framework (section 2.3) was applied for the rasterization of polygons. Cells present in 2017 but absent in 2024 were classified as abandoned terraced areas, while cells absent in 2017 but present in 2024 were classified as expanded terraced areas. Abandoned terraced areas were subsequently mapped to tabulate their crop types and land suitability classes. A random sample of 100 agriculture plot polygons was used to check if the expanded terraced areas were recently (2017–2024) constructed by earthwork machinery or revitalised abandoned terraces, using Google Earth imagery, Orthophotos from 2014 and 2019 (available through the Department of Land and Surveys) and field visits.

## 3. Results

### 3.1. Selection of DEM resolution and study area classification

The total agricultural area above 500 m elevation, as derived from the original crop plot polygons was 10,640.7 ha. Rasterization using different DEM resolutions resulted in an underrepresentation, with the 5-m and 30-m DEMs retaining 99.8% and 88.3% of the original polygons area, respectively (Table 3). This reduction is due to the omission of small and fragmented parts of the polygons at coarser resolution. The 5-m DEM preserves finer spatial detail, captures small and fragmented agriculture, and produces a rasterized plot area that is closer to the original polygon-based area, thus justifying its selection for the LSA.

**Table 3**
Fractions (%) of agricultural areas above 500 m elevation, classified by broader crop type and slope class (≥15% and < 15%), based on 5-m and 30-m DEMs using majority area rule.

| Crop classes | Delineated using 5-m elevation model | | Delineated using 30-m elevation model | |
|---|---|---|---|---|
| | Area ≥ 15% slope | Area < 15% slope | Area ≥ 15% slope | Area < 15% slope |
| Fallow (%) | 6.2 | 5.2 | 5.3 | 5.7 |
| Grapes (%) | 18.9 | 13.1 | 19.2 | 14.7 |
| Irrigated annual (%) | 1.8 | 0.9 | 1.2 | 1.0 |
| Irrigated perennial (%) | 10.9 | 5.2 | 9.3 | 5.4 |
| Olive & Carob (%) | 6.4 | 3.5 | 5.0 | 4.0 |
| Rainfed annual (%) | 9.9 | 12.3 | 8.8 | 15.1 |
| Pasture (%) | 2.8 | 2.0 | 3.9 | 1.5 |
| Not subsidised (%) | 0.3 | 0.2 | 0.3 | 0.1 |
| Total % | 57.5 | 42.5 | 52.5 | 47.5 |
| Total area (ha) | 10,626.7 | | 9,396.4 | |

The crop types recorded by CAPO for individual plots, aggregated into eight broader categories (Table 3) revealed that agricultural land accounts for only 3.8% of the total study area (elevation > 500 m and slope ≥ 15%), of which a minor portion (~329.4 ha) is pasture or non-subsidised. Table 4 summarizes the classification of study area based on TP (all agriculture areas except pasture and non-subsidised) and TA (all non-agriculture areas and land classified as pasture and non-subsidised).

### 3.2. Correlation analysis

Correlation analysis (Fig. 3) revealed a strong positive relationship between elevation and precipitation (ρ = 0.73) and a strong negative relationship between elevation and temperature (ρ = –0.91). Precipitation and temperature were also strongly negatively correlated (ρ = –0.72). Due to these high correlations (|ρ| > 0.7), precipitation and temperature were excluded, resulting in a final set of 14 covariates. These remaining covariates displayed only low to moderate correlations (|ρ| < 0.7) and hence were retained as predictors for ML simulations.

### 3.3. Pu-learning predictions on unlabelled examples

The ElkanotoPuClassifier estimated the labelling frequency "c" as 0.72 based on a 20% hold-out subset of labelled positive examples. The fraction of labelled positives among all positives (ρ′), derived from the full training set was estimated as 0.77. Under the SCAR assumption, the classifier was evaluated on 20% of unseen data using PU modified metrics (Table 2). This PU-learning model achieved a Recall of 84.6% indicating that most true positive examples in the unlabelled dataset were correctly identified. The Precision of 81.5% reflected a low rate of false-positive predictions. The resulting F1 score of 83% confirms a strong balance between Precision and Recall. These metrics demonstrate that the trained model generalizes effectively to unseen data. Notably, 88.3% of the unlabelled examples had a predicted probability (to be positive) of less than 0.1 (Table 5), highlighting that the vast majority of unlabelled instances are highly unlikely to be positive and can be used to compute a reliable set of negative examples.

### 3.4. Selection of reliable negatives and suitability mapping

Under a constant sample size for training and validation, the classifier trained exclusively on positive examples and bin with predicted probabilities below 0.1 (i.e., Set-1) achieved the highest performance. Incorporating data from higher probability bins led to a decline in classification performance, suggesting the presence of noise. Specifically, positive examples incorrectly treated as negatives in the training sets. Table 5 summarizes the reduction in Precision, Recall, F1-score and AUC when including data from higher probability bins and evaluated using random-split, identifying 0.1 as a reliable threshold for confidently selecting negative examples (Set-1). The performance of Set-1 for the suitable class was further evaluated in Table S2.3, reporting the Omission Error, Commission Error, and Figure of Merit. A similar trend was observed under spatial-CV, where Set-1 again yielded the best performance, with an average Precision of 0.78, Recall of 0.86, F1-score of 0.82, and AUC of 0.88 across five repetitions. The slightly lower performance of spatial-CV compared to the random split suggests some uncertainty in model performance attributable to spatial autocorrelation. Performance metrics for each repetition, along with average values across all three sets obtained from the spatial-CV, are reported in Table S2.2 in Supplementary Material 2.

**Table 4**
Classification of the study area (elevation ≥ 500 m and slope ≥ 15%), based on Terrace-Present (TP) and Terrace-Absent (TA) land.

| Grid type | Grid class | Grid area (ha) |
|---|---|---|
| Presence of terraced land (positive, labelled examples) | TP | 5,778.47 |
| Absence of terraced land (unlabelled examples) | TA | 155,238.11 |
| Total study area | – | 161,016.58 |

The supervised machine learning model resulted in a continuous probability distribution of land suitability. The *k*-means analysis indicated that the rate of reduction in SSE notably decreased for more than four clusters, suggesting that utilizing four clusters provide a suitable balance between capturing variation and maintaining interpretability. The class thresholds identified by the *k*-means clustering of the four classes were subsequently used to categorize the continuous suitability values into four distinct classes (Table 6). The classification revealed that only 7% of the total study area is "Highly Suitable" for agricultural terraces (Fig. 4). Within this highly suitable zone, 37% is already under existing terrace cultivation. Notably, a substantial potential area of approximately 7,000 ha remains available within the highest suitability class, presenting significant opportunities for future terrace development. To support decision-making, the developed suitability map and associated data layers have been made accessible through Google Earth Engine application and presented to land managers and government official (https://ee-ameena.projects.earthengine.app/view/land-suitability-analysis).

### 3.5. Feature importance analysis

The results of the SHAP feature importance analysis are presented using the mean |SHAP| plot (Fig. 5) and the beeswarm plot (Fig. 6). The mean |SHAP| plot presents the overall influence of each feature on terrace suitability, ranking them according to their average contribution to model predictions. The beeswarm plot illustrates the distribution of SHAP values across all grid cells, with each point representing the SHAP value of a continuous covariate for an individual grid. It shows how high and low values of continuous covariates influence terrace suitability, driving predictions in either positive or negative direction.

The feature importance analysis identified land cover, obtained from the CORINE (Table 1), as the most influential parameter for terrace suitability, accounting for 23.9% of the total mean |SHAP| value. Coniferous forests, which occupy a substantial portion of the Troodos Mountains (53.9%), were found to exert a negative influence on terrace suitability (Fig. 7). In contrast, vineyards, fruit and berry plantations, and agricultural areas with significant natural vegetation collectively represent 10.7% of the study area yet contain 58.2% of the terraced land, thereby demonstrating a positive influence on terrace suitability (Fig. 7). The areal distribution of all feature categories over the four suitability classes are presented in Tables S1.1–S1.16 in the Supplementary Material.

Slope was identified as the second most influential parameter, accounting for 15.0% of the total mean |SHAP| value. A negative linear relationship between slope and suitability was observed (Fig. 8a). Although terraces are generally recommended for slopes greater than or equal to 15%, areas with slopes exceeding 50% exhibited negative SHAP values, indicating unsuitability for terrace construction on very steep terrain (Fig. 8a).

Tree cover density emerged as the third most influential feature, contributing 14.9% to the total mean |SHAP| value. Areas with lower tree density were classified as more suitable for terrace construction. Specifically, grids without any tree cover occupied 90% of the highly suitable area. Nonetheless, in rare instances, a minor fraction of grid cells with higher tree cover also contributed positively to the model output (Fig. 8b). An intersection of these high tree density grids with land cover revealed that they are predominantly associated with specific land-cover categories i.e., Fruit trees and berry plantations; Complex cultivation patterns; and agriculture land with significant natural vegetation. As illustrated in Fig. 7, these land-cover categories exert a strong positive influence on terrace suitability predictions. Since land cover is the most influential predictor in the model, its strong positive contribution of these categories (mean SHAP value is + 1.6) dominates

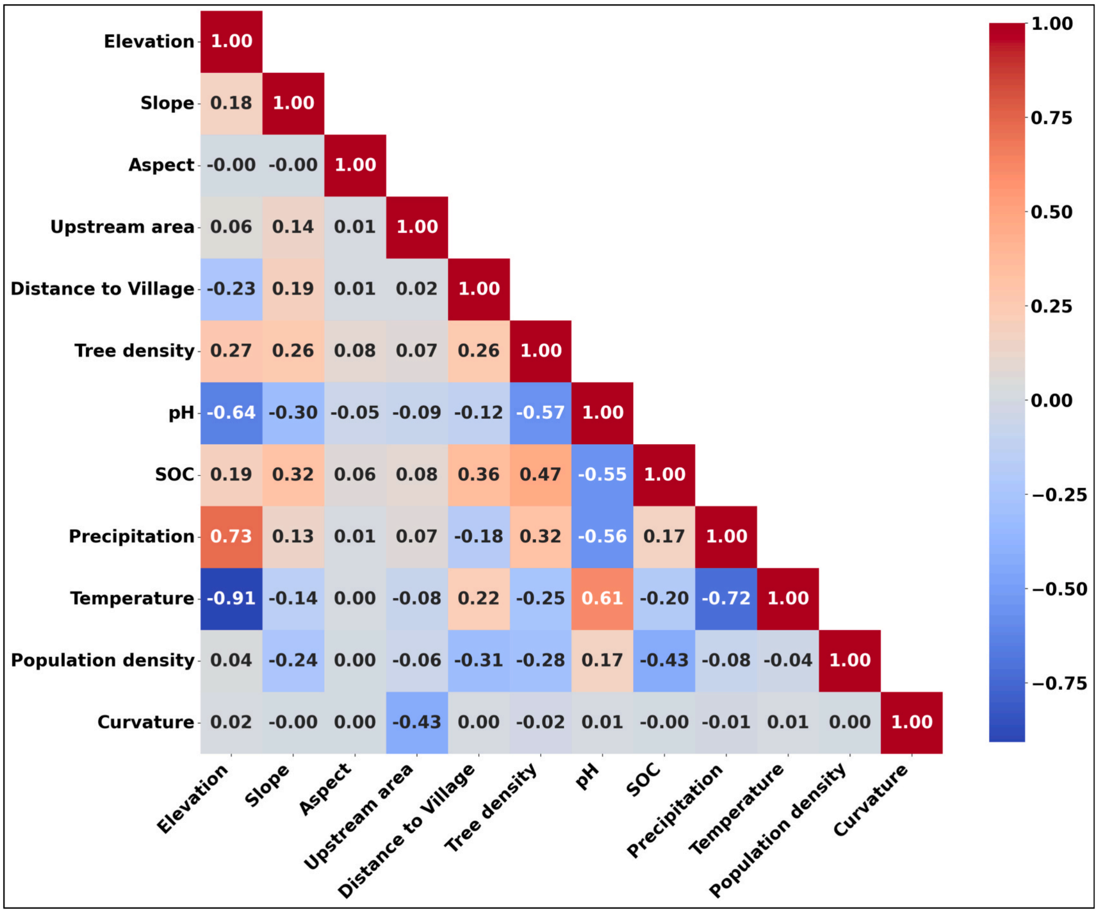


**Fig. 3.** Spearman's rank correlation (ρ) heatmap for continuous covariates identified for terrace suitability assessment.

**Table 5**
Fraction of examples in bins with low probability of being positive in total dataset, number of examples in three training sets (with 1,951,251 positive examples in each set) and corresponding XGBoost performance (random-split) used to identify the most reliable negative set.

| Probability bin | 0–0.1 | 0.1–0.2 | 0.2–0.3 | Performance | | | |
|---|---|---|---|---|---|---|---|
| Share in total dataset (%) | 88.3 | 3.0 | 1.6 | Precision | Recall | F1 | AUC |
| Set-1 | 45,579,651 | – | – | 0.83 | 0.88 | 0.86 | 0.93 |
| Set-2 | 44,137,365 | 1,442,286 | – | 0.74 | 0.83 | 0.78 | 0.87 |
| Set-3 | 43,427,212 | 1,419,080 | 733,359 | 0.68 | 0.80 | 0.74 | 0.82 |

**Table 6**
Classification of study area and TP grid cells based on suitability classes.

| Suitability classes | *k*-means threshold | Total area (ha) | Area under TP cells (ha) |
|---|---|---|---|
| Not suitable | 0.001–0.112 | 129,136.58 | 22.77 |
| Marginally suitable | 0.112–0.370 | 13,174.85 | 363.51 |
| Moderately suitable | 0.370–0.696 | 7,566.97 | 1,250.33 |
| Highly suitable | 0.696–0.989 | 11,138.18 | 4,141.86 |

the general negative influence of high tree cover density (mean SHAP value is −0.40), resulting in an overall net positive impact on suitability for these minor fractions of grid cells.

The planning zones of Cyprus were identified as the fourth most influential contributor, with a total contribution of 9.7% to the suitability prediction. The majority of the study area (58.1%), including 28.4% of TP cells falls within protected zones designated for areas of natural beauty, forests, landscapes, rivers and archaeological sites. Within these zones, only a limited fraction, ranging from 1% to 8% of a cadastral plot, is permitted for building construction, with a maximum building height restriction of two floors. For instance, Zone 'Z3′ covers 44.6% of the total protected land in the study area, with building construction permitted on only 1% of each cadastral plot and limited to a single story. Agricultural and livestock zones cover 27.7% of the study area, including 58.7% of TP cells. Notably, 63.3% of the highly suitable land lies within agricultural and livestock zones, while 24.6% is located within protected zones.

Demographically, communities with larger populations tend to be associated with higher suitability for terrace agriculture. For instance, Omodos community covers a total area of 13.85 km$^2$ and has 334 inhabitants, with 48.8% of its area classified as highly suitable for terrace agriculture. In contrast, Mylikouri, despite being one of the largest communities in terms of area (85.2 km$^2$), has only 11 inhabitants, with 95.3% of its land covered by coniferous forest and almost no land suitable for terrace agriculture. Additionally, areas of the Troodos Mountains located within 2 km of the centre of a settlement account for the majority of the highly suitable land (65.4%). Suitability declines sharply with increasing settlement distance, and almost no land is classified as

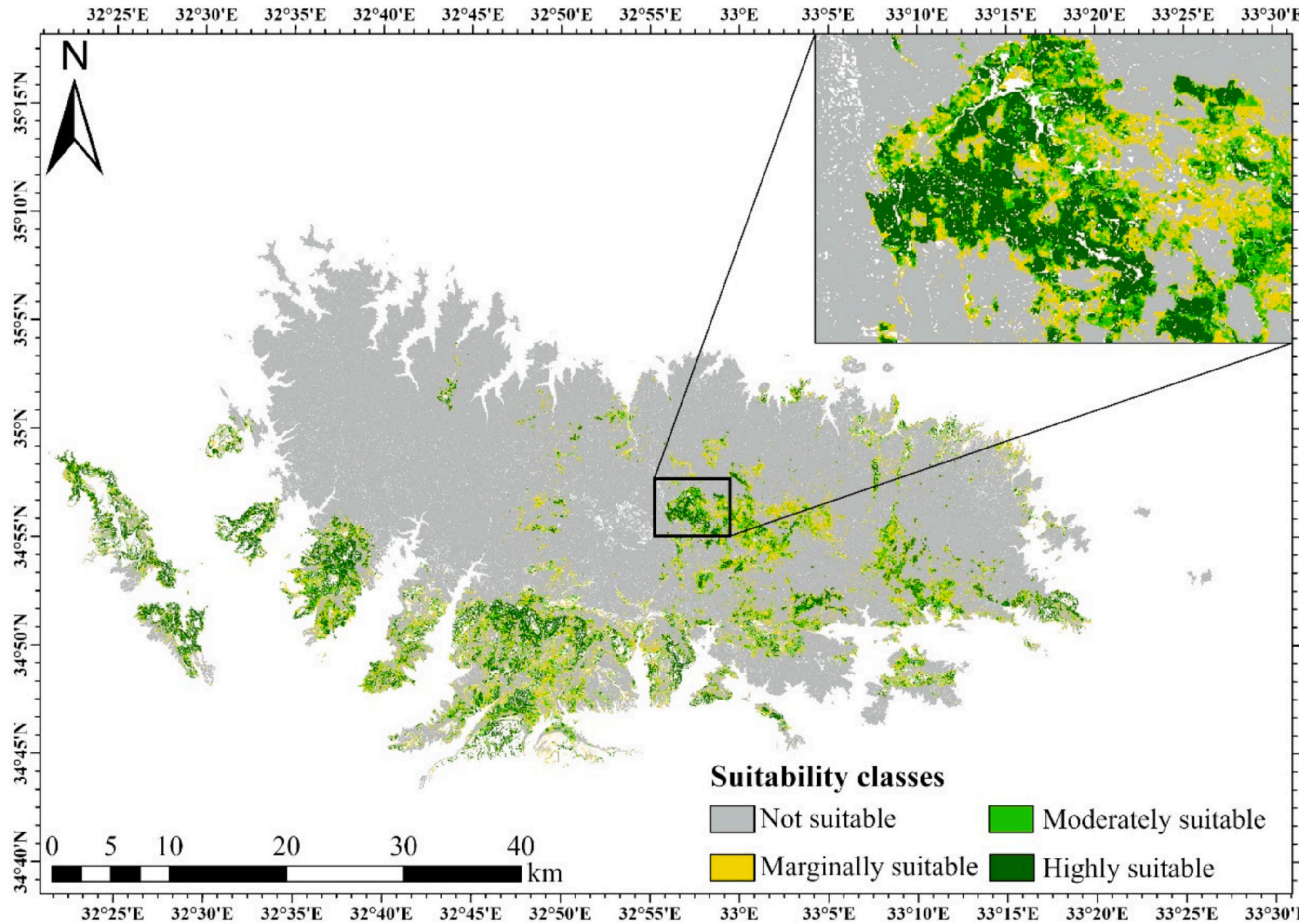


**Fig. 4.** Terrace suitability map for Troodos Mountains, Cyprus.

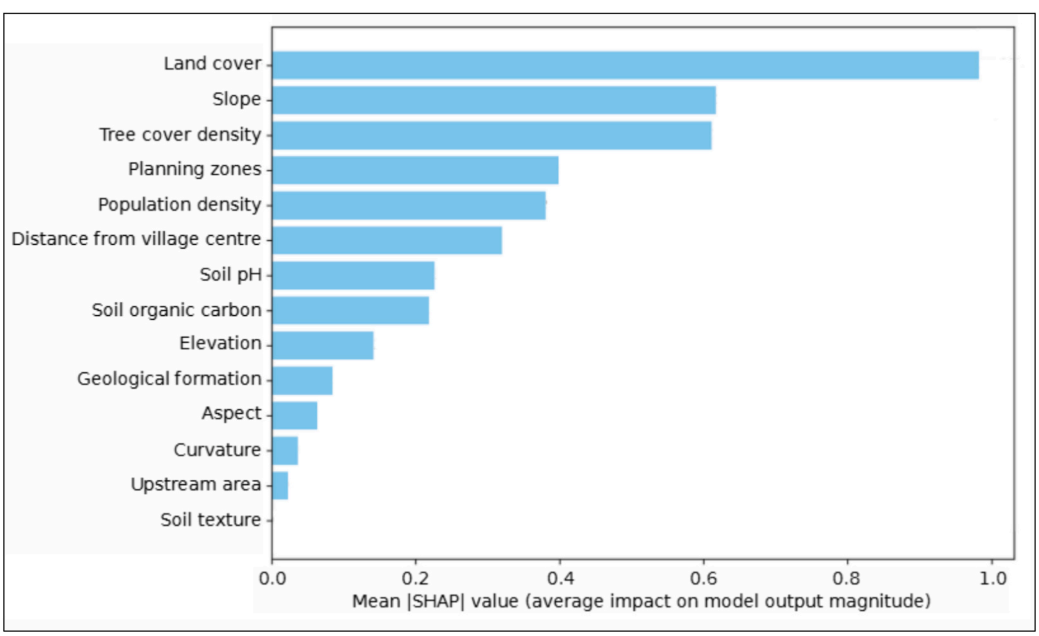


**Fig. 5.** Mean |SHAP| enabling global interpretation of feature contributions to model predictions for individual features.

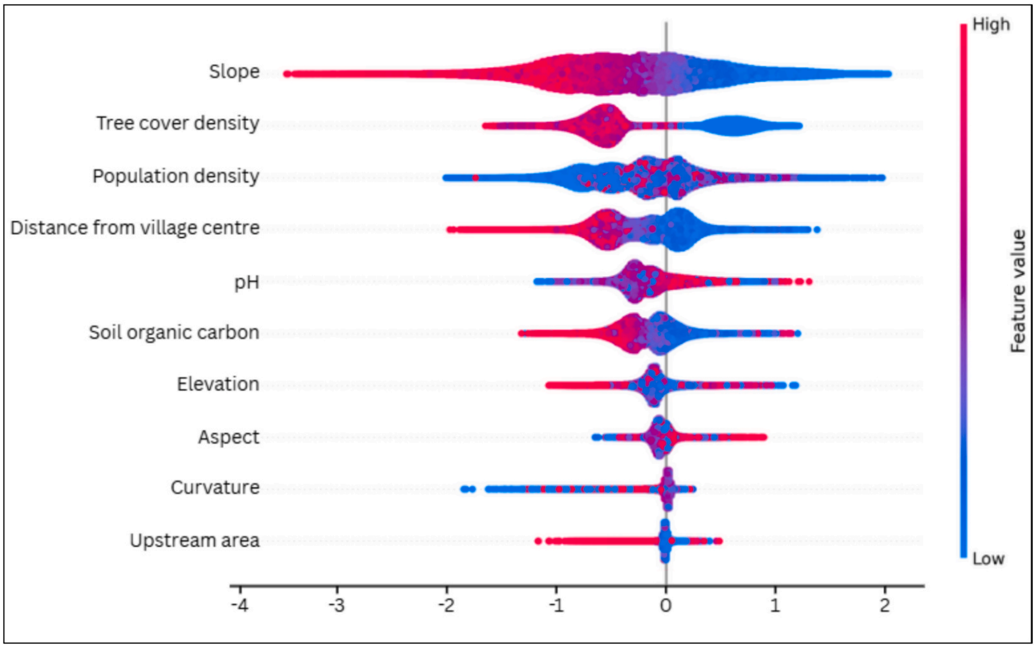


**Fig. 6.** SHAP beeswarm plot illustrating the positive and negative contributions of low and high values of continuous covariates to terrace suitability.

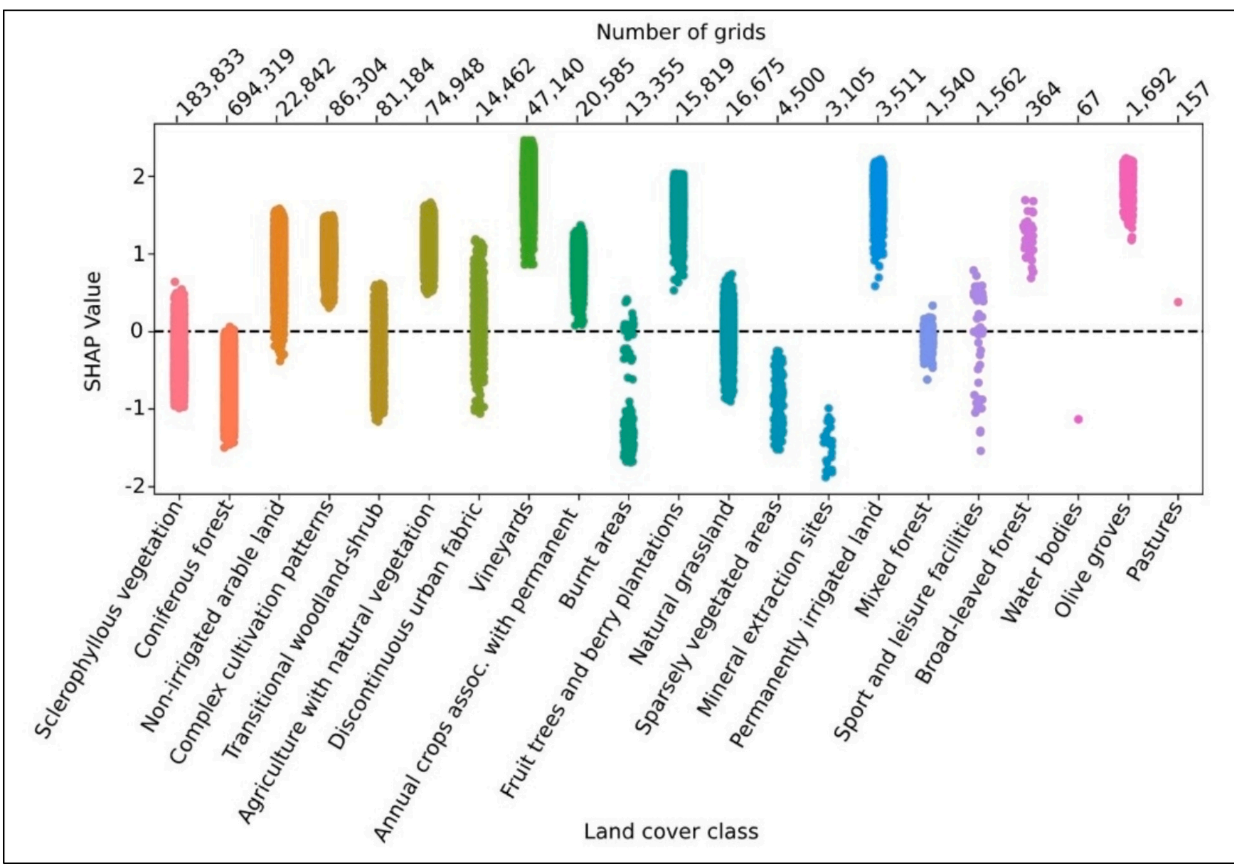


**Fig. 7.** SHAP dependence plot for land cover, highlighting the contribution of individual land cover class to model predictions.

suitable beyond 8 km.

Soil properties demonstrated a moderate contribution to the suitability prediction, with soil pH and soil organic carbon contributing 5.5% and 5.3%, respectively, to the total mean |SHAP| value. The majority (82%) of the highly suitable areas have a pH of 7–8, and 92% have soil organic carbon values of 2–3%. Soil texture showed no measurable influence on suitability, as 98.5% of the study area is occupied by medium texture soil. The geological map showed more differentiation. The majority of highly suitable areas for terrace agriculture are found in the calcareous Pakhna (35.8%) and Lefkara (18.5%) formations, which

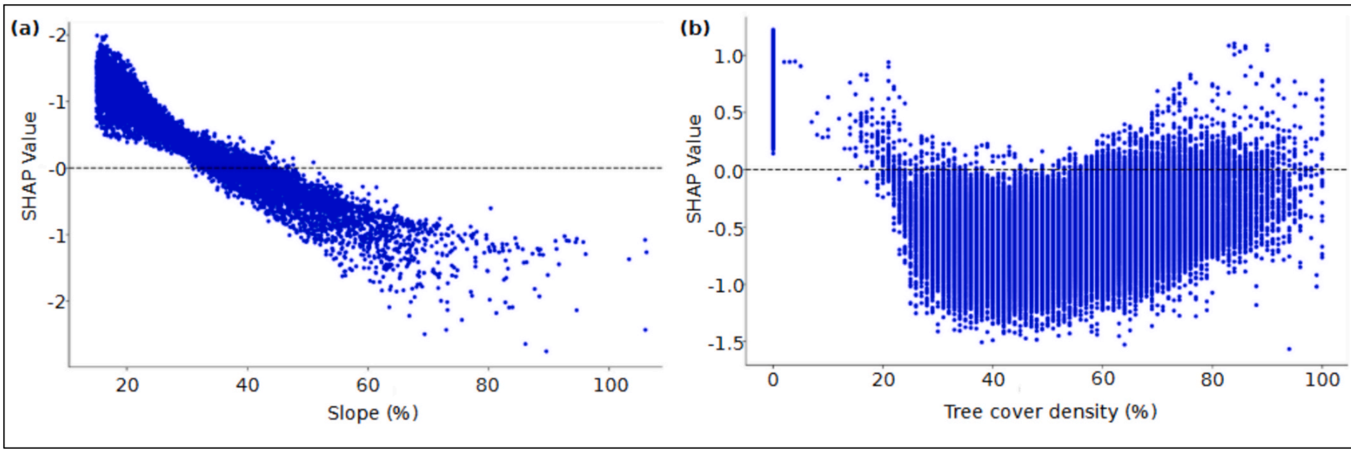


**Fig. 8.** SHAP dependence plot for terrain slope (left) and tree cover density (right), highlighting the contribution of these covariates to the model predictions.

occupy 15.5% of the study area. This is followed by the ophiolite diabase dykes (12.8%), and gabbro (9.2%) formations. Although diabase formations dominate the Troodos Mountains, covering half (50.5%) of the study area, most of this area is occupied by coniferous forest.

Topographically, 60% of the highly suitable areas are located at elevations of 500–700 m, which cover 43.6% of the study area. The climatic parameters, which are strongly correlated with elevation and not used in the suitability analysis, showed that 80% of the highly suitable land was found in rainfall zones with an average annual rainfall of 500–700 mm and a mean annual temperature range of 15–18 °C. Aspect and curvature have no significant influence on suitability, with highly suitable areas almost equally distributed between convex (47.9%) and concave (46.7%) surfaces, and slightly more located on southeast (15.2%), south (14.5%), and east (14.1%) slopes compared to other directions (10–13%). Additionally, areas with smaller upstream catchment areas are identified as more suitable for terrace construction.

The exclusion of the CORINE land cover covariate resulted in minor variations in the suitability predictions. For instance, the areas classified as 'not suitable' and 'highly suitable' decreased by 3.0% and 2.1%, respectively. Tree cover density emerged as the most influential contributor to the suitability predictions (accounting for 17.5% of the total mean |SHAP|), followed closely by slope (17.0% of the total mean |SHAP|). The results of this second analysis are provided in Supplementary Material 2. Tables S2.4 and S2.5 present a comparative overview of the areal distribution of the study area, agricultural crop plots of 2024, and CORINE land cover across suitability classes, with and without the land use covariate. Figs. S2.2 and S2.3 present the results of feature importance obtained by second analysis.

### *3.6. Recent land use dynamics: Abandonment vs expansion of mountain terraced land*

The comparison of agricultural crop plots between 2017 and 2024 revealed 12% (828.0 ha) terrace expansion and 29% (2,026.3 ha) terrace abandonment, with respect to the total terraced area in 2017. More than half of the abandoned terraces are located within highly and moderately suitable zones. Furthermore, over one-quarter of the abandonment area corresponds to vineyards, of which 10.8% falls within the highly suitable class (Table 7). Nevertheless, grapes still remain the largest crop class, in terms of area, in 2024 (Table 3). Field and image-based inspection of the 100 randomly selected agriculture plots that were identified as expansion areas indicated that 18% of these plots were recently (2017–2024) constructed by earthwork machinery, generally reshaping previously abandoned terraced areas to improve access and field operations. This evaluation also found that two out of the 50 field-inspected plots were not terraced, although neighbouring plots were, indicating minor uncertainties in the dataset.

**Table 7**
Crop type and suitability class fractions of mountain terrace land abandoned between 2017 and 2024.

| Crop class | Not suitable | Marginally suitable | Moderately suitable | Highly suitable |
|---|---|---|---|---|
| Pasture (%) | 0.5 | 0.8 | 0.7 | 1.0 |
| Fallow (%) | 2.2 | 3.3 | 3.2 | 5.7 |
| Grapes (%) | 4.3 | 5.6 | 5.5 | 10.8 |
| Irrigated annual (%) | 0.4 | 0.6 | 0.7 | 0.8 |
| Irrigated perennial (%) | 4.8 | 5.5 | 4.6 | 4.9 |
| Not subsidized (%) | 2.8 | 1.2 | 0.6 | 0.8 |
| Olive & Carob (%) | 2.8 | 3.1 | 2.8 | 3.4 |
| Rainfed annual (%) | 0.6 | 1.2 | 1.4 | 4.7 |
| Rainfed perennial (%) | 2.9 | 2.2 | 1.6 | 2.1 |
| Total (%) | 21.3 | 23.5 | 21.1 | 34.3 |
| Total abandoned area (ha) | 2,026.3 | | | |

## 4. Discussion

This study demonstrates the potential of PU-learning for LSA, addressing the common challenges of limited positive samples or contaminated negative samples. Although PU learning is a robust approach for addressing the lack of known negative data, it poses challenges for model evaluation. The absence of negative instances limits the applicability of standard predictive performance metrics, which require true class labels for all instances. Saunders and Freitas (2022), in their review of 51 PU-learning studies, noted a widespread lack of separate reporting of Precision and Recall, emphasizing the need to present both independently. The current study utilized PU-learning metrics modified under the SCAR (Tabatabaei et al., 2021) to calculate Precision and Recall. In this study, the developed PU classifier achieved a Recall of 84.6%, Precision of 81.5% and F1 score of 83%, underscoring its strong predictive performance. These performance metrics are higher than those reported in recent PU-learning applications for GIS mapping. For instance, (Zhang et al., 2024) evaluated different ML models with and without PU-learning for landslide susceptibility assessment and reported that the best-performing model (an RF classifier combined with the Elkan and Noto PU-learning framework), validated under the SCAR assumption, achieved an average

Recall of 56.7%. Similarly, Wang et al. (2024) reported a Precision of 57.07% for the developed PU classifier used for winter wheat mapping.

The LSA demonstrated that land cover, tree density and regulatory constraints in mountain environments play a decisive role in determining terrace suitability. These constraints are intended to safeguard forest and mature trees from anthropogenic disturbances. According to the FAO (2015), forest land in Cyprus is defined as areas exceeding 0.5 ha with trees taller than 5 m and a canopy cover greater than 10%, excluding land predominantly used for agricultural or urban purposes. Protection criteria also follow ecological metrics, considering species-specific growth patterns rather than fixed dimensions. For instance, a pine may require > 100 cm trunk diameter (at breast height) for protection, while a golden oak or juniper may qualify at > 50 cm (Department of Forests, Ministry of Agriculture, Natural Resources and Environment, 2012) . Exceptions occur in harsh environments where old trees have limited size. A recent study by Zenonos et al. (2025) found a strong correlation between tree cover density and tree counts for three regions of Cyprus, with a correlation coefficient of 0.93 for Troodos Park, which overlaps the area of the current study. Although, the results of the current study aligned with these regulatory restrictions by identifying coniferous forests as not suitable, and areas with no tree covers as highly suitable, some anomalies persisted. For instance, 5.4% of the TP cells falls within coniferous forests. This is likely linked to the limitations of CORINE land cover map used in this study. This dataset, developed by European Environment Agency through computer-assisted photointerpretation of high-resolution imagery, applies a minimum mapping unit of 25 ha and a minimum feature width of 100 m (Aune-Lundberg and Strand, 2021). This generalization can mask small-scale heterogeneity, causing terrace patches to be aggregated into surrounding classes, with orchards or fragmented land class often misclassified as forest or sclerophyllous areas (Aune-Lundberg and Strand, 2021) . Although the CLCplus Backbone dataset provides higher spatial resolution (10 m) and a smaller minimum mapping unit (0.5 ha) (European Environment Agency, 2022), it does not distinguish between agricultural trees (orchards) and forest trees. Similar classification challenges in mountainous regions have been reported by Li et al. (2025), who emphasized that rugged terrain, complex illumination, and highly interwoven land cover types increase the likelihood of confusion between forests, grasslands, and terraces. Considering these challenges, future applications of the presented ML framework could benefit from incorporating high-resolution canopy datasets capable of differentiating horticultural from natural forest trees and distinguishing mature forest trees that should be protected from new terrace developments.

Feature importance analysis identified slope (15.0% of total mean |SHAP|) as the second most influential factor and elevation (3.4% of total mean |SHAP|) as a moderate contributor to terrace suitability. These findings align with those of Li et al. (2025), where a 10-m global terrace distribution map was developed using Sentinel-2 imagery and topographic features derived from a 30-m DEM. Their study employed advanced deep learning methods in combination with cloud computing platforms for extraction of terrace pixels from images. They performed feature importance analysis using a RF classifier and identified slope as critical contributor (37.4%) and elevation as moderate contributor (9.4%) for terrace mapping. The SHAP analysis of the current study also emphasizes the need to consider maximum slope thresholds for terrace suitability. Specifically, areas with slopes exceeding 50% (43.0% of study area) were identified as unsuitable for terracing in this region. For instance, on the southern slopes of the Troodos Mountains, presence of ridges and valleys create steep terrain, causing several abandoned terraces to be classified as unsuitable due to slope gradients exceeding 50%, even when surrounded by highly suitable land and TP cells. Comparable thresholds for slope limitations have been reported in previous studies. For instance, Mesfin (2016) identified a maximum slope threshold of 60% for the construction of bench terraces on deep and stable soils, while recommending a reduced limit of 30% for unstable soils in Ethiopia. Similarly, Sabir (2021) documented slope ranges of 40–60% for terraced terrain in the Atlas of Morocco. FAO (1988) also suggested a maximum threshold of 47% for the construction of bench terraces. Together, these findings highlight that the slope threshold identified for the Troodos Mountains aligns closely with these reported ranges.

Numerous studies have reported a positive association between terrace presence and high population density (e.g., Brown et al., 2025; Leceta et al., 2024). A similar relationship was also observed in the Troodos Mountains. Although, almost half of the in the study area have less than 100 inhabitants, communities with a population of more than 100 inhabitants occupied 75% of highly suitable areas. The absence of dense population in mountain regions reduces the socio-economic incentives for terrace construction and maintenance, contrasting with regions in Asia (Chen et al., 2022). Additionally, previous studies on terrace systems relative to their distance from settlement areas indicated that abandonment rates and degradation tend to increase with increase in distance from settlements, largely due to higher maintenance costs, reduced accessibility, and less frequent cultivation (Kizos et al., 2009; Long et al., 2024). Consistent with these studies, finding of the current research shows that areas located close to the settlements are highly suitable. Suitability declines sharply with increasing distance from these settlement areas. These findings indicate that terraces in the Troodos Mountains are predominantly maintained by a small number of mountain families residing in nearby villages, reflecting a pattern consistent with global trends of mountain agriculture (Zoumides et al., 2022; Wymann von Dach et al., 2014).

## 5. Conclusion

Terraced agriculture in Mediterranean mountains provides vital ecosystem services but is increasingly threatened by climate change and socio-economic pressures. To advance sustainable terrace agriculture in Mediterranean mountain environments, this study presents an integrated predictive modelling framework combining GIS with data-driven ML technique for terrace suitability assessment. Unlike traditional supervised approaches that require fully labelled datasets, the proposed PU-learning framework identified suitable areas for terrace agriculture using only positive and unlabelled data. The challenges associated with the validation of PU-learning classifiers were addressed using the SCAR assumption and PU-modified metrics. The proposed framework was applied to develop a terrace suitability map for the Troodos Mountains in Cyprus, achieving a Precision of 81.5% and Recall of 84.6%. These scores are higher than those of recently published PU-learning mapping applications, although direct comparison is limited by differences in study area, data structure, class definition and validation protocol. The developed land suitability map classified 7% of the total study area as highly suitable for terracing, of which 2.6% is under existing terrace cultivation, leaving a substantial area for possible expansion. This research demonstrates that the proposed framework can effectively model terrace suitability, identify opportunities for terrace expansion, and provide a replicable methodology for mountain landscapes.

Feature importance analysis identified land cover as the most influential predictor of terrace suitability and tree cover density as the third. Slope emerged as the second most important predictor with an upper threshold of 50%, closely aligning with limits reported for terraces in the literature. Planning zones, population density, settlement proximity, soil properties and geological formations demonstrated a moderate influence on terrace suitability. The comparative analysis of the terraced agriculture area in 2017 and 2024 revealed 29% abandonment and 12% revitalisation. The observed abandonment aligns with previously reported patterns driven by socioeconomic shifts and labour constraints, while the revitalisation, which was partially driven by mechanised construction, reflects the renewed interest in mountain agriculture. These contrasting trends highlight the need for informed planning and guidelines to support mountain agriculture.

To strengthen the methodological basis of data-driven land suitability assessment, future studies could investigate different PU-learning algorithms, as well as hybrid approaches that combine supervised classification with PU-learning (for example, RF and SVM models wrapped within the Elkan and Noto framework) and compare their performance for different types of datasets. Efforts could also focus on developing and integrating reliable land-use and soil datasets to better understand their influence on terrace suitability. The use of high-resolution Unmanned Aerial Vehicle based surveys could further enhance the identification of terraces, while reducing uncertainties. Additionally, terrace suitability mapping could be combined with ecosystem service assessments, such as soil conservation, carbon storage and biodiversity habitats, to support more holistic and sustainable land management decisions.

## CRediT authorship contribution statement

**Aman Kumar Meena:** Writing – original draft, Software, Methodology, Data curation. **Christos Zoumides:** Writing – review & editing, Conceptualization. **Hakan Djuma:** Writing – review & editing, Conceptualization. **Ioannis Sofokleous:** Writing – review & editing, Data curation. **Corrado Camera:** Writing – review & editing. **Adriana Bruggeman:** Writing – review & editing, Supervision, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors acknowledge the High-Performance Computing Facility at the Cyprus Institute for support with computational and storage resources, and the Cyprus Agricultural and Payment Organization for providing the data essential to this study. This research was supported by the REACT4MED Project (GA 2122), which is funded by PRIMA, the Partnership for Research and Innovation in the Mediterranean Area, a Programme supported by Horizon 2020, the European Union's Framework Programme for Research and Innovation.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2026.105454.

## Data availability

The rasterized dataset and python scripts have been made available on Zenodo at: https://doi.org/10.5281/zenodo.18978123.

## References

Alonso, N., Viedma-Guiard, A., Simón-Rojo, M., Córdoba-Hernández, R., 2025. Agricultural land suitability analysis for land use planning. The case of the Madrid Region. Land 14, 134. https://doi.org/10.3390/land14010134.

Aune-Lundberg, L., Strand, G.-H., 2021. The content and accuracy of the CORINE Land Cover dataset for Norway. Int. J. Appl. Earth Obs. Geoinf. 96, 102266. https://doi.org/10.1016/j.jag.2020.102266.

Batunacun, W.R., Lakes, T., Nendel, C., 2021. Using Shapley additive explanations to interpret extreme gradient boosting predictions of grassland degradation in Xilingol, China. Geosci. Model Dev. 14 (3), 1493–1510. https://doi.org/10.5194/gmd-14-1493-2021.

Bekker, J., Davis, J., 2020. Learning from positive and unlabelled data. A Survey. Mach. Learn. 109, 719–760. https://doi.org/10.1007/s10994-020-05877-5.

Bi, B., Li, J., Luo, T., Wang, B., Yang, C., Shen, L., 2025. Positive–unlabelled learning-based hybrid models and interpretability for groundwater potential mapping in karst areas. Water 17 (10), 1422. https://doi.org/10.3390/w17101422.

Brown, A.G., Pears, B., Cucchiaro, S., Tarolli, P., Lang, A., Zhao, P., Walsh, K., Van Ost, K., Albert, R.-M., Alonso-Eguiluz, M., Vokotopoulos, L., Tsartsidou, G., Molinari, A., Stagno, A., Ghislandi, S., Wei, W., Fallu, D., 2025. The geoarchaeology of agricultural terraces in Europe. construction, resilience and implications for sediment delivery. Geoarchaeology 40, e70008. https://doi.org/10.1002/gea.70008.

Camera, C., Bruggeman, A., Hadjinicolaou, P., Pashiardis, S., Lange, M.A., 2014. Evaluation of interpolation techniques for the creation of gridded daily precipitation ($1 \times 1$ km$^2$); Cyprus, 1980–2010. J. Geophys. Res.-Atmos. 119, 693–712. https://doi.org/10.1002/2013JD020611.

Camera, C., Djuma, H., Bruggeman, A., Zoumides, C., Eliades, M., Charalambous, K., Abate, D., Faka, M., 2018. Quantifying the effectiveness of mountain terraces on soil erosion protection with sediment traps and dry-stone wall laser scans; Troodos Mountains, Cyprus. Catena 171, 251–264. https://doi.org/10.1016/j.catena.2018.07.017.

Carranza-García, M., García-Gutiérrez, J., Riquelme, J.C., 2019. A framework for evaluating land use and land cover classification using convolutional neural networks. Remote Sens. 11 (3), 274. https://doi.org/10.3390/rs11030274.

Chen, D., Wei, W., Chen, L., 2017. Effects of terracing practices on water erosion control in China. A meta-analysis. Earth-Sci. Rev. 173, 109–121. https://doi.org/10.1016/j.earscirev.2017.08.007.

Chen, Q., Wen, Y., Zhang, X., Zhu, Z., 2022. Evolutionary overview of terrace research based on bibliometric analysis in Web of Science from 1991 to 2020. Int. J. Environ. Res. Public Health 19, 7796. https://doi.org/10.3390/ijerph19137796.

Christofi, C., Bruggeman, A., Kuells, C., Constantinou, C., 2020. Hydrochemical evolution of groundwater in gabbro of the Troodos Fractured Aquifer. A comprehensive approach. Appl. Geochem. 114, 104524. https://doi.org/10.1016/j.apgeochem.2020.104524.

Dax, T., Schroll, K., Machold, I., Derszniak-Noirjean, M., Schuh, B., Gaupp-Berghausen, M., 2021. Land abandonment in mountain areas of the EU. An inevitable side effect of farming modernization and neglected threat to sustainable land use. Land 10, 591. https://doi.org/10.3390/land10060591.

Deng, C., Zhang, G., Liu, Y., Nie, X., Li, Z., Liu, J., Zhu, D., 2021. Advantages and disadvantages of terracing. A comprehensive review. Int. Soil Water Conserv. Res. 9 (3), 344–359. https://doi.org/10.1016/j.iswcr.2021.03.002.

Department of Forests, Ministry of Agriculture, Natural Resources and Environment, 2012. Tree nature monuments of Cyprus. Press and Information Office, Nicosia, Cyprus.

Desloires, J., Ienco, D., Botrel, A., Ranc, N., 2022. Positive unlabelled learning for satellite images' time series analysis. An application to cereal and forest mapping. Remote Sens. 14 (1), 140. https://doi.org/10.3390/rs14010140.

Djuma, H., Bruggeman, A., Camera, C., Zoumides, C., 2017. Combining qualitative and quantitative methods for soil erosion assessments: an application in a sloping Mediterranean watershed, Cyprus. Land Degrad. Develop. 28, 243–254. https://doi.org/10.1002/ldr.2571.

Elkan, C., Noto, K., 2008. Learning classifiers from only positive and unlabelled data, in: Li, Y., Liu, B., Sarawagi, S. (Eds.), Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 213–220. https://doi.org/10.1145/1401890.1401920.

European Environment Agency, 2022. Copernicus Land Monitoring Service – CLC+ Backbone production, including raster and vector products based on satellite input data from 2017/2018/2019: CLC+ Backbone product specification and user manual (Issue 3.0). Framework Service Contract No. EEA/DIS/R0/19/012.

European Parliament and Council of the European Union, 2013. Regulation (EU) No 1305/2013 on support for rural development by the European Agricultural Fund for Rural Development (EAFRD) and repealing Council Regulation (EC) No 1698/2005). Off. J. Eur. Union L 347, 487–548.

FAO, 1988. Watershed management field manual. Slope treatment measures and practices (FAO Conservation Guide No. 13/3). Food and Agriculture Organization of the United Nations, Rome. https://www.fao.org/4/ad083e/ad083e07.htm.

FAO, 2015. Global Forest Resources Assessment 2015: Country report, Cyprus. Food and Agriculture Organization of the United Nations, Rome. https://openknowledge.fao.org/handle/20.500.14283/az195e.

FAO, 2019. Mountain agriculture. Opportunities for harnessing Zero Hunger in Asia. Bangkok.

Franco, W., Barbera, F., Bartolucci, L., Felizia, T., Focanti, F., 2020. Developing intermediate machines for high-land agriculture. Dev. Eng. 5, 100050. https://doi.org/10.1016/j.deveng.2020.100050.

Ganesan, M., Andavar, S., Raj, R.S.P., 2021. Prediction of land suitability for crop cultivation using classification techniques. Braz. Arch. Bio. Technol. 64. https://doi.org/10.1590/1678-4324-2021200483.

Georgiades, P., Economou, T., Proestos, Y., Araya, J., Lelieveld, J., Neira, M., 2025. Global projections of heat stress at high temporal resolution using machine learning. Earth Syst. Sci. Data 17 (3), 1153–1171. https://doi.org/10.5194/essd-17-1153-2025.

Geng, S., Shi, P., Zong, N., Zhu, W., 2019. Agricultural land suitability of production space in the Taihang Mountains, China. Chin. Geogr. Sci. 29 (5). https://doi.org/10.1007/s11769-019-1075-6.

Karongo, J., Mwaniki, J.I., Ndiritu, J., et al., 2025. Sorghum yield prediction based on remote sensing and machine learning in conflict affected South Sudan. Sci. Rep. 15, 4469. https://doi.org/10.1038/s41598-025-89030-z.

Kizos, T., Dalaka, A., Petanidou, T., 2009. Farmers' attitudes and landscape change. Evidence from the abandonment of terraced cultivations on Lesvos. Greece. Agric. Hum. Values 27 (2), 199–212. https://doi.org/10.1007/s10460-009-9206-9.

Kyriakou, M.S., 2022. Digital elevation model of Cyprus 5m (2014). Zenodo. https://doi.org/10.5281/zenodo.6006664.

Lasanta, T., Arnaez, J., Ruiz-Flaño, P., Lana-Renault, N., Arnáez, J., 2013. Agricultural terraces in the Spanish mountains. An abandoned landscape and a potential resource. Bol. Asoc. Geógr. Esp. 63, 487–491.

Leceta, F., Binder, C., Mader, C., Mächtle, B., Marsh, E., Dietrich, L., Reindel, M., Eitel, B., Meister, J., 2024. The impact of agriculture on tropical mountain soils in the western Peruvian Andes. A pedo-geoarchaeological study of terrace agricultural systems in the Laramate region (14.5° S). Soil 10 (2), 727–761. https://doi.org/10.5194/soil-10-727-2024.

Lekshmi, G.S., Aryadevi Remanidevi, D., Pushpalatha, R., Gangadharan, B., Hariprasad, K.M., 2025. Enhancing coconut yield potential. A climate-smart land suitability analysis using machine learning. Smart Agric. Technol. 12, 101087. https://doi.org/10.1016/j.atech.2025.101087.

Li, W., Guo, Q., Elkan, C., 2011. Can we model the probability of presence of species without absence data? Ecography 34 (6), 1096–1105. https://doi.org/10.1111/j.1600-0587.2011.06888.x.

Li, W., Liu, Y., Liu, Z., Gao, Z., Huang, H., Huang, W., 2022. A positive-unlabelled learning algorithm for urban flood susceptibility modeling. Land 11 (11), 1971. https://doi.org/10.3390/land11111971.

Li, Y., Tian, F., Zhang, M., Zeng, H., Ahmed, S., Qin, X., Liu, Y., Wang, L., Fan, R., Wu, B., 2025. A 10-meter global terrace mapping using Sentinel-2 imagery and topographic features with deep learning methods and cloud computing platform support. Int. J. Appl. Earth Obs. Geoinf. 139, 104528. https://doi.org/10.1016/j.jag.2025.104528.

Long, M., Zhao, Y., Zhou, C., Li, X., Su, L., Zou, Y., 2024. Analysis of factors influencing terrace abandonment based on unmanned aerial photography and farmer surveys. A case study in Jianhe, Guizhou. Land Degrad. Dev. 35 (2), 757–771. https://doi-org.proxy2.library.illinois.edu/10.1002/ldr.4950.

Luan, W., Li, G., Zhong, B., Geng, J., Li, X., Li, H., He, S., 2023. Improving dryland urban land cover classification accuracy using a classical convolution neural network. Land 12 (8), 1616. https://doi.org/10.3390/land12081616.

Lundberg, S. M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. (NeurIPS 2017). arXiv.1705.07874. https://doi.org/10.48550/arXiv.1705.07874.

Maharjan, S.K., Sterck, F.J., Raes, N., Poorter, L., 2021. Temperature and soils predict the distribution of plant species along the Himalayan elevational gradient. J. Trop. Ecol. https://doi.org/10.1017/S026646742100050X.

Mali, Y., Rathod, V.U., Kulkarni, M.M.S., Mokal, P., Patil, S., Dhamdhere, V., Birari, D.R., 2023. A comparative analysis of machine learning models for soil health prediction and crop selection. Int. J. Intell. Syst. Appl. Eng. 11 (10s), 811–828. https://www.ijisae.org/index.php/IJISAE/article/view/4479.

Ministry of Agriculture, Natural Resources and Environment, 2021. Cyprus Rural Development Programme 2014–2020 / Πρόγραμμα Αγροτικής Ανάπτυξης 2014–2020 (CCI 2014CY06RDNP001, Version 8.1).

Moeys, J., 2024. The Soil Texture Wizard. R functions for plotting, classifying, transforming and exploring soil texture data.

Mesfin, A., 2016. A field guideline on bench terraces design and construction. Addis Ababa.

Negussie, K.G., Gebrekidan, B.H., Wyss, D., Kappas, M., 2024. Assessing land suitability for leguminous crops in the Okavango River Basin. A multicriteria and machine learning approach. Int. J. Appl. Earth Obs. Geoinf. 135, 104284. https://doi.org/10.1016/j.jag.2024.104284.

Pawłuszek-Filipiak, K., Lewandowski, T., 2025. The impact of feature selection on XGBoost performance in landslide susceptibility mapping using an extended set of features: a case study from Southern Poland. Appl. Sci. 15 (16), 8955. https://doi.org/10.3390/app15168955.

Poggio, L., de Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. SOIL 7, 217–240. https://doi.org/10.5194/soil-7-217-2021.

Radočaj, D., Jurišić, M., 2022. GIS-based cropland suitability prediction using machine learning. A novel approach to sustainable agricultural production. Agronomy 12 (9), 2210. https://doi.org/10.3390/agronomy12092210.

Ramos, M.C., Cots-Folch, R., Martínez-Casasnovas, J.A., 2007. Sustainability of modern land terracing for vineyard plantation in a Mediterranean mountain environment – the case of the Priorat region (NE Spain). Geomorphology 86 (1–2), 1–11. https://doi.org/10.1016/j.geomorph.2006.08.004.

Romeo, R., Manuelli, S. R., Geringer, M., Barchiesi, V., (Eds.), 2021. Mountain farming systems – Seeds for the future. Sustainable agricultural practices for resilient mountain livelihoods. FAO, Rome. https://doi.org/10.4060/cb534.

Sabir, M., 2021. The terraces of the Anti-Atlas: From abandonment to the risk of degradation of a landscape heritage. Water 13, 510. https://doi.org/10.3390/w13040510.

Sahu, N., Das, P., Saini, A., Varun, A., Mallick, S.K., Nayan, R., Aggarwal, S.P., Pani, B., Kesharwani, R., Kumar, A., 2023. Analysis of tea plantation suitability using geostatistical and machine learning techniques. A case of Darjeeling Himalaya, India. Sustainability 15 (13), 10101. https://doi.org/10.3390/su151310101.

Saunders, J.D., Freitas, A.A., 2022. Evaluating the predictive performance of positive-unlabelled classifiers. A Brief Critical Review and Practical Recommendations for Improvement. Arxiv Preprint arXiv.

Saunders, J.D., Freitas, A.A., 2025. Automated machine learning for positive-unlabelled learning. Appl. Intell. 55, 875. https://doi.org/10.1007/s10489-025-06706-9.

Shawon, S.M., Ema, F.B., Mahi, A.K., Niha, F.L., Zubair, H.T., 2025. Crop yield prediction using machine learning. An extensive and systematic literature review. Smart Agric. Technol. 10, 100718. https://doi.org/10.1016/j.atech.2024.100718.

Sinchana, G., Choodarathnakara, A., Arpitha, G., 2024. Machine learning models to develop land suitability map for coffee cultivation by integrating CHIRPS and SRTM DEM. Research Square. .

Singh, R., Behera, M.D., Das, P., Rizvi, J., Dhyani, S.K., Biradar, C.M., 2022. Agroforestry suitability for planning site-specific interventions using machine learning approaches. Sustainability 14 (9), 5189. https://doi.org/10.3390/su14095189.

Sofokleous, I., Bruggeman, A., Michaelides, S., Hadjinicolaou, P., Zittis, G., Camera, C., 2021. Comprehensive methodology for the evaluation of high-resolution WRF multi-physics precipitation simulations for small, topographically complex domains. J. Hydrometeorol. 22. https://doi.org/10.1175/JHM-D-20-0110.1.

Stanchi, S., Freppaz, M., Agnelli, A., Reinsch, T., Zanini, E., 2012. Properties, best management practices and conservation of terraced soils in Southern Europe (from Mediterranean areas to the Alps). A review. Quat. Int. 265, 90–100. https://doi.org/10.1016/j.quaint.2011.09.015.

Sun, K., Hu, Y., Lakhanpal, G., Zhou, R.Z., 2023. Spatial cross-validation for GeoAI, in: handbook of geospatial artificial intelligence, chapter 10. GeoAI Lab. Department of Geography. University at Buffalo, Stevenson High School.

Tabatabaei, S. A., Klein, J., Hoogendoorn, M., 2021. Estimating the F1 score for learning from positive and unlabelled examples. In LOD 2020. Machine Learning, Optimization, and Data Science pp. 1–12. https://doi.org/10.1007/978-3-030-64583-0_15.

Taghizadeh-Mehrjardi, R., Nabiollahi, K., Rasoli, L., Kerry, R., Scholten, T., 2020. Land suitability assessment and agricultural production sustainability using machine learning models. Agronomy 10 (4), 579. https://doi.org/10.3390/agronomy10040579.

Tarolli, P., 2018. Agricultural terraces special issue preface. Land Degrad. Dev. 29 (10), 3544–3548. https://doi.org/10.1002/ldr.3129.

Tarwidi, D., Pudjaprasetya, S.R., Adytia, D., Apri, M., 2023. An optimized XGBoost-based machine learning method for predicting wave run-up on a sloping beach. MethodsX 10, 102119. https://doi.org/10.1016/j.mex.2023.102119.

Torgbor, B.A., Rahman, M.M., Brinkhoff, J., Sinha, P., Robson, A., 2023. Integrating remote sensing and weather variables for mango yield prediction using a machine learning approach. Remote Sens. 15, 3075. https://doi.org/10.3390/rs15123075.

Wang, H., Yu, F., Xie, J., Wan, H., Zheng, H., 2024. Mapping winter wheat using ensemble-based positive unlabeled learning approach. Photogramm. Eng. Remote Sens. 90, 483–491. https://doi.org/10.14358/PERS.23-00038R2.

Wei, T., Shi, F., Wang, H., Tu, W.-W., Li, Y.-F., 2020. MIXPUL. Consistency-based augmentation for positive and unlabelled learning. arXiv preprint. https://doi.org/10.48550/arXiv.2004.09388.

Wymann von Dach, S., Romeo, R., Vita, A., Wurzinger, M., 2013. Mountain Farming IS Family Farming. A Contribution from Mountain Areas to the International Year of Family Farming 2014. FAO, CDE, and BOKU, Rome, Italy. http://www.fao.org/docrep/019/i3480e/i3480e.pdf.

Xing, W., Zhou, C., Li, J., Wang, W., He, J., Tu, Y., Cao, X., Zhang, Y., 2022. Suitability evaluation of tea cultivation using machine learning technique at town and village scales. Agronomy 12, 2010. https://doi.org/10.3390/agronomy12092010.

Yang, P., Li, X.-L., Mei, J.-P., Kwoh, C.-K., Ng, S.-K., 2012. Positive-unlabelled learning for disease gene identification. Bioinformatics 28 (20), 2640–2647. https://doi.org/10.1093/bioinformatics/bts504.

Yuan, C., Yang, H., 2019. Research on K-Value selection Method of K-Means Clustering Algorithm. J 2 (2), 226–235. https://doi.org/10.3390/j2020016.

Zenonos, A., Li, S., Brandt, M., Sciare, J., Ciais, P., 2025. AI-powered estimation of tree covered area and number of trees over the Mediterranean island of Cyprus. Front. Remote Sens. 6, 1498217. https://doi.org/10.3389/frsen.2025.1498217.

Zhang, D., Jindal, D., Roy, N., Vangla, P., Frost, J.D., 2024. Enhancing landslide susceptibility mapping using a positive-unlabelled machine learning approach. A case study in Chamoli, India. Geoenviron. Disasters 11, 21. https://doi.org/10.1186/s40677-024-00281-w.

Zhao, L., Ma, H., Dong, J., Wu, X., Xu, H., Niu, R., 2023. A comparative study of landslide susceptibility mapping using bagging PU learning in class-prior probability shift datasets. Remote Sens. 15, 5547. https://doi.org/10.3390/rs15235547.

Zheng, Y., Peng, H., Zhang, X., Zhao, Z., Gao, X., Li, J., 2019. DDI-PULearn. a positive-unlabelled learning method for large-scale prediction of drug-drug interactions. BMC Bioinform. 20 (Suppl 19), 661. https://doi.org/10.1186/s12859-019-3214-6.

Zoumides, C., Bruggeman, A., Giannakis, E., Camera, C., Djuma, H., Eliades, M., Charalambous, K., 2017. Community-based rehabilitation of mountain terraces in Cyprus. Land Degrad. Develop. 28 (1), 95–105. https://doi.org/10.1002/ldr.2586.

Zoumides, C., Bruggeman, A., Giannakis, E., Kyriakou, N., 2022. A future for mountain terraces. Experiences from Mediterranean wineries. Mt. Res. Dev. 42 (3). https://doi.org/10.1659/mrd-journal-d-21-00031.1.

Zoumides, C., Bruggeman, A., Djuma, H., 2025. Revitalising terraced landscapes: Co-production of ecosystem services for sustainable futures. Geogr. Sustain. 6, 100306. https://doi.org/10.1016/j.geosus.2025.100306.