Article

Assessing Soil and Land Suitability of an Olive–Maize Agroforestry System Using Machine Learning Algorithms

Asif Hayat [1], Javed Iqbal [1], Amanda J. Ashworth [2,*] and Phillip R. Owens [3]

[1] Institute of Geographical Information Systems, School of Civil and Environmental Engineering, National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan; ahayatdir@gmail.com (A.H.); javed@igis.nust.edu.pk (J.I.)
[2] Poultry Production and Product Safety Research Unit, USDA-Agricultural Research Service, 1260 W. Maple St, Fayetteville, AR 72701, USA
[3] Dale Bumpers Small Farms Research Center, USDA-Agricultural Research Service, 6883 S. Hwy 23, Booneville, AR 72927, USA; phillip.owens@usda.gov
* Correspondence: amanda.ashworth@usda.gov

Abstract: Exponential population increases are threatening food security, particularly in mountainous areas. One potential solution is dual-use intercropped agroforestry systems such as olive (*Olea europaea*)–maize (*Zea mays*), which may mitigate risk by providing multiple market sources (oil and grain) for smallholder producers. Several studies have conducted integrated agroforestry land suitability analyses; however, few studies have used machine learning (ML) algorithms to evaluate multiple variables (i.e., soil physicochemical properties and climatic and topographic data) for the selection of suitable rainfed sites in mountainous terrain systems. The goal of this study is therefore to identify suitable land classes for an integrated olive–maize agroforestry system based on the Food and Agriculture Organization (FAO) land suitability assessment framework for 1757 km^2 in Khyber Pakhtunkhwa province, Pakistan. Information on soil physical and chemical properties was obtained from 701 soil samples, along with climatic and topographic data. After determination of land suitability classes for an integrated olive–maize-crop agroforestry system, the region was then mapped through ML algorithms using random forest (RF) and support vector machine (SVM), as well as using traditional techniques of weighted overlay (WOL). Land suitability classes predicted by ML techniques varied greatly. For example, the S1 area (highly suitable) classified through RF was 9%↑ than that of SVM, and 8%↓ than that through WOL. The area of S2 (moderately suitable) classified through RF was 18%↑ than that of SWM and was 17%↓ than the area classified through WOL; similarly, the S3 (marginally suitable) class area via RF was 27%↓ than that of SVM, and 45%↓ than the area classified through WOL. Conversely, the area of N2 (permanently not suitable class) classified through RF and SVM was 6%↑ than the area classified through WOL. Model performance was assessed through overall accuracy and Kappa Index and indicated that RF performed better than SVM and WOL. Crop suitability limitations of the study area included high elevation, slope, pH, and large gravel content. Results can be used for sustainable intensification in mountainous rainfed regions by expanding intercrop agroforestry systems in developing nations to close yield gaps.

Keywords: random forest; support vector machine; intercropping agroforestry; land suitability

1. Introduction

Due to rapid population increase, demand for food has also risen nationally. Similarly, rising income in developed countries has increased demands for high-value food products [1]. Additionally, demand for edible olive oil in the U.S. has increased three-fold over the last 20 years [2]. The population of Pakistan has substantially increased over the past 20 years along with demands for edible oil, as Pakistan is the fourth largest edible oil importer [3]. Agriculture worldwide is a major contributor to deforestation and land degradation [4]; however, agroforestry systems may protect and mitigate soil losses, improve soil

health, provide ecosystem services, and improve global carbon sequestration [5]. In addition, intercropping agroforestry systems have major potential to maximize food production per unit area, provide ecosystem services, and improve economic viability. Agroforestry is a collective name for land-use systems and practices in which woody perennials are deliberately integrated with crops and/or animals on the same land-management unit. Such practices may attain more sustainable land use and can enhance production per unit area, while improving the livelihood of rural economies [6]. Maize is one of the prominent cereal crops and is utilized for food, fodder, feed, and other value-added products, with 46% of maize being produced under rainfed conditions [7]. Therefore, to enhance the production of rainfed systems, combat population growth rises by 3% from the last two decades, and address concurrent food security issues in Pakistan, researchers set out to evaluate land suitability for an integrated olive- and maize-crop agroforestry system.

Every region has varying agroforestry systems which depend on crop suitability and climatic factors [8–13]. In the District of Lower Dir Pakistan, 54.4% is rainfed and 50.5% of farmers follow agro-silvo-pastoral systems where tree species (*Ailanthus altissima*, *Morus alba*, *Juglans regia*, and *Populous nigra*) are integrated with grain and fodder crops to enhance ecosystem services [14]. However, little work has investigated the suitability of maize–olive intercrop agroforestry systems in Pakistan for closing yield gaps and improving food security. Production of agricultural lands in developing countries is very low compared to developed countries [15], and to overcome this issue, proper land suitability analyses are needed to close yield gaps and improve soil health. Such indices have a multitude of functions, spanning land-use planning to matching crops and soil for optimization, and soil conservation. The land assessment framework's goals aim to identify negative impacts and gains of agricultural lands and there is now a strong focus on environmental impacts, as well as on benefits from ecological and environmental resources [16]. Further, land suitability analysis is the first step towards agricultural land-use planning for testing suitability of a particular application quantitatively [17]. The qualitative evolution comprises data on climatic, hydrological, topographic, vegetative, and edaphic characteristics of the study area, with the quantitative evolution of land suitability focusing on the productivity of soils [18]. The FAO land evolution framework and *Land Suitability*, *part III: crop requirements* [19,20] are widely used for land suitability analysis of integrated annual crops and perennial tree species.

Land suitability mapping offers knowledge to land managers and policymakers to make important decisions for reducing soil erosion and assessing sustainable land use. There is a deficiency of land suitability maps in Pakistan because land suitability surveys and mapping have not adopted modern approaches [21–24]. Moreover, increasing demand for high-resolution soil maps around the world is being met by digital soil mapping as a method for generating spatial soil information. In developing countries, digital soil mapping is becoming increasingly important as field surveys are less frequent, time-consuming, and costly, and soil surveys of individual federation units can no longer be afforded, but can be integrated with agroforestry-crop suitability mapping [25].

Machine learning (ML) models can strategically combine data from multiple sources and identify relationships by learning from these massive datasets [26]. These ML models can play a vital role in creating a relationship between soil data and ancillary data, which help to understand the spatial and temporal changes during land suitability classification based on soil, climatic, and auxiliary variables [18,27]. Artificial neural networks, partial least squares regression, support vector machine (SVM), generalized additive models, genetic programming, regression tree models, k nearest neighbor regression, adaptive neuro-fuzzy inference scheme, and random forests (RF) are some machine learning models that are used for establishing such relations [18,28–32]. Among these machine learning models, RF and SVM are the most widely used techniques for digital soil mapping because of their high accuracy, ease of use, and robustness. Ancillary data for crop suitability mapping are most widely obtained from digital elevation models, geo-spatial data, and remote sensing data [18,32–34].

ML models have been extensively used to build digital soil maps in the past few years [35], although few researchers use these ML models for mapping land suitability in agroforestry systems [14,36]. As an example, Chen et al. [37] used classification and regression tree algorithm to predict a suitable tree species for afforestation in Jinping County, Guizhou Province of China. Similarly, Taghizadeh-Mehrjardi et al. [18] assessed optimum land suitability for a rainfed wheat and barley system in the Kurdistan province, western Iran using ML algorithms from data obtained from 100 soil samples along with climatic and topographic variables. Jiang et al. [38] used k-means clustering to more precisely establish and manage agroforestry systems based on digital soil mapping and tree data (e.g., absolute growth rate and diameter at breast height). Kharel et al. [39] used ML to tease apart factors that affect system-level output (i.e., plant production (tree and forage), soil factors, and animal response based on grazing preference) and found that RF models helped identify system drivers in integrated tree–livestock systems. However, no studies to date have tested and evaluated ML models [i.e., RF, SVM, and traditional techniques of weighted overlay (WOL)] for their ability to identify suitable land classes for an integrated olive and maize agroforestry system for 1757 km^2 in Khyber Pakhtunkhwa province, Pakistan. Therefore, the goal of this study is to identify suitable land classes for an integrated olive and maize agroforestry system based on the FAO land suitability assessment framework, in an effort to close yield gaps.

2. Materials and Methods

2.1. Description of Experimental Region

The study was conducted throughout the district Dir Lower (34.9161° N, 71.8097° E) of Malakand division in the Khyber Pakhtunkhwa province of Pakistan at an altitude of 567 m from sea level. It has a total area of about 1757 km^2 (Figure 1) and the geographical boundary of the district is connected to Dir upper in the northeast and north, with district Swat to the east, district Malakand to the south, district Bajaur to the southwest, and Afghanistan to the west through the Durand line. Climatic conditions of the study area are characterized by relatively medium to high annual rainfall (1000–1200 mm). In the winter, temperature is reported as low as −6 °C min and as high as 38 °C, and in summer it ranges from 15 to 40 °C. The land-use landcover of the study area consists of 30% agricultural land, 15% barren land, 29% forests, 14% rivers, and 12% urban area.
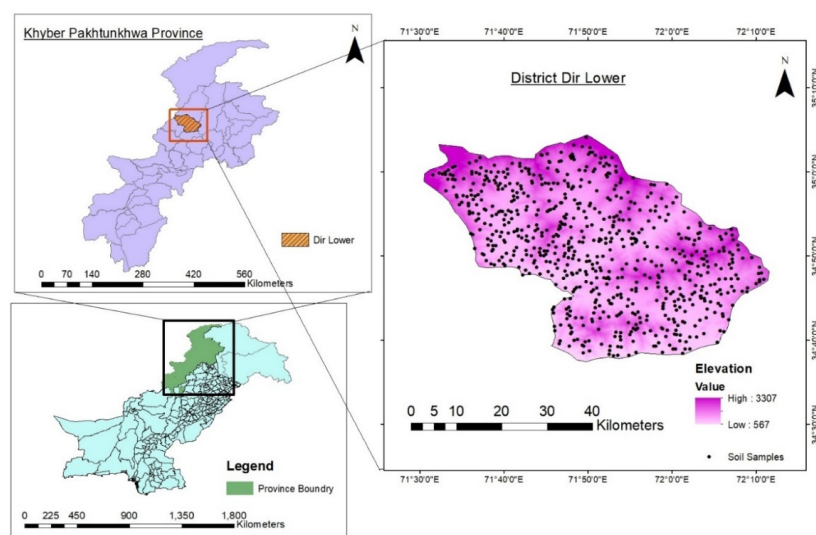
Figure 1. Study area along with soil sample (*n* = 701) locations in Khyber Pakhtunkhwa province.

2.2. Soil Sampling and Lab Analysis

A random soil sampling approach was used to ensure accurate representation of soil properties in the study area [40]. The total number of soil samples collected throughout the

study area was 701 (dependent variables). Soil samples were collected with a tube auger (Figure 1), and the spatial position of each soil sample was recorded through a handheld Global Positioning System (GPS) device. The climatic data primarily included annual temperature and rainfall. The topographic data were retrieved from DEM, and indices like NDVI, SAVI, and BI were generated from satellite imagery, as shown in Table 1. The analysis for this study included multiple steps, with a flowchart on methodology provided in Figure 2.

Table 1. Data types and sources used in ML and traditional techniques.

Data Type	Source
Soil data	Random point sampling from the study area
Climate data	World Bank Climate Change Knowledge Portal
Topographic data	Digital elevation model (DEM)
Indices	Satellite imagery of sentinel 2A
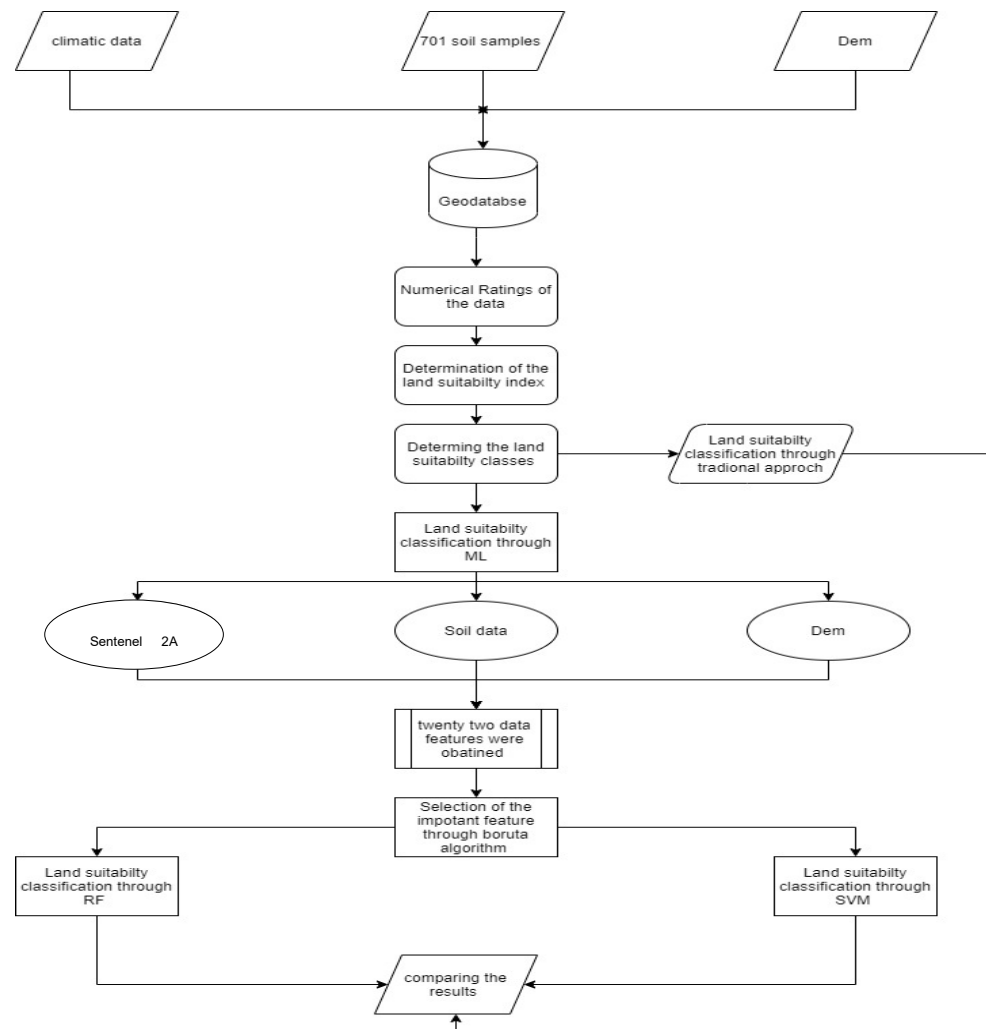
Figure 2. Flow chart methodology represents the land suitability classification through machine learning (RF, SVM) and traditional technique (WOL).

Soil samples were placed in bags and sent to the laboratory for further examination. Soil samples were dried, crushed, and then sieved through a 2 mm sieve [41]. Soil physical properties such as soil texture were identified through the hydrometric method [42]. Chemical properties, including soil organic carbon, were determined through wet combustion in

which the evolved CO_2 was passed through a carrier stream in which potassium iodide and silver sulfate (KI, $Ag_2 SO_4$) were present, which prevents Cl interference. To trap traces of water vapor H_2SO_4, granular Zn and anhydride were used after CO_2 was weighed as absorbed by Mikohbite in a Nesbitt bulb. In calcareous soil, CO_3 was removed by a pretreatment to determine the organic carbon. Then, soil EC and pH were measured potentiometrically in a 2:1 water-volume-to-soil mass [43–46].

2.3. Land Suitability Analysis

2.3.1. Determining Land Suitability Index Using Parametric Methods

Following determination of land suitability classes for an olive–maize agroforestry system, classification of land suitability classes through the weighted overlay technique was done by stacking all layers of climatic, soil, and topographic data, and then weights were assigned. Thereafter, selection of important features with ancillary and collected data for the land suitability classes was performed through a Boruta RF algorithm [47]. A map of land suitability classes was then created using ArcGIS PRO 2.5.0 to compare results of WOL, RF, and SVM classification. Soil pH, organic carbon, soil texture, and slope along with climatic data (temperature and rainfall) were used for developing land suitability indices and then scored [48]. Soil EC and $CaCO_3$ had no growth limitation for olive and maize crops, so those data were excluded during land suitability classification.

The square root technique was used to calculate land suitability indices for a maize–olive agroforestry system after choosing variables. This approach first assesses the climate. As a result, the climatic index is computed based on a rating of climate features. This climatic index was transformed to a climate rating, which was then combined with the soil and topographic ratings to create the final land suitability index. Computing the land suitability index employs the square root approach in the following Equation (1).

$$I_L = Rmin \sqrt{\frac{A}{100} + \frac{B}{100} + \frac{C}{100}} \tag{1}$$

I_L is the land index, Rmin is the minimal rating, and the other rating values were *A*, *B*, and *C*. The parametric technique assigns a value from 0 to 100 to each soil property depending on the limitation imposed by each soil feature [49].

2.3.2. Determination of Land Suitability Classes

Land suitability classes were identified after computing a land index for an agroforestry system comprising olive and maize crops. There were five land suitability classes: highly suitable (S1), moderately suitable (S2), marginally suitable (S3), and non-suitable (N1 and N2) [19]. The land indices for the S1, S2, S3, N1, and N2 classes were 85–100, 60–85, 40–60, 25–40, and 0–25, respectively [48]. The non-suitable (N) land class was thought to have significant restrictions that could never, or only slightly, be addressed by improvement techniques.

2.4. Construction of Models for the Classification of the Land Suitability Classes

2.4.1. Traditional Method of Weighted Overlay (WOL)

The land suitability classification for an olive–maize agroforestry system was mapped through a traditional method of weighted overlay and then analyzed with machine learning algorithms (RF and SVM) in ArcGIS PRO 2.5.0 [50]. After determining land suitability classes, each data layer was overlayed through WOL analysis based on its relative influence [22,51]. WOL is one of the most common techniques used for the determination of suitability classes. The weights of each feature were determined through the AHP model (analytic hierarchy process) [52].

2.4.2. Machine Learning Models: Random Forest (RF) and Support Vector Machine (SVM)

Random forest is a machine learning algorithm using multiple decision/classification trees. The main idea is to utilize the bootstrap sampling technique to choose *k* samples

at random from the original dataset, which are then brought back in as a new training set. The classification trees are then built to build a random forest of *k* decision tree, and the affiliation of new samples is selected by majority vote based on the findings of each classification tree. The algorithm's aim is to strengthen decision trees by integrating numerous decision trees, with each tree being created based on an autonomously generated sample set.

To evaluate land suitability classes for olive–maize agroforestry systems, two ML algorithms were employed on a set of auxiliary variables including terrain characteristics and satellite imagery in conjunction with climatic data. The final ML model was built on the training data, and the testing data (point data) were used to control the construction process and select the optimal parameters. Because of their successful applicability in previous research [18,33,34,53], we chose two machine learning models: RF and SVM. Both RF and SVM function effectively when data are not available on a large scale. In this work, the extensive linkages between the predicted classes of land suitability and auxiliary variables that may be deduced by the exploration of relevant auxiliary variables are noteworthy.

The basic principle behind SVM is to map data to a high-dimensional space in order to explore the ideal classification hyperplane by reducing the upper limit of classification error. SVM focuses on determining the separating hyperplane $\omega^T x + b$ for linearly separable data. Equation (2) describes its objective function.

$$\text{argmax}_{w,b} \left\{ \min_n \left(\text{label} \cdot \left(\omega^T x + b \right) \right) \cdot \frac{1}{||w||} \right\} = \pi r2 \tag{2}$$

where ω is the normal vector that specifies the hyperplane's direction, label is the category label, and label $(\omega^T x + b)$ is the constraint. Equation (2) may be modified into Equation (3) by using Lagrange multipliers.

$$\left\{ \max\alpha \left(\sum_{i=1}^{m} \alpha - \frac{1}{2} m \sum_{ij-1}^{m} \text{label}^{(i)} \cdot \text{label}^{(i)} \cdot \alpha_i \alpha_j < x^{(i)}, x^{(j)} > \right) \right\} C \geq \alpha \geq 0, \, m \sum_{i=1}^{m} \text{label}^{(i)} = 0 \tag{3}$$

where α is the displacement that determines the distance of the hyperplane from the origin and C is the relaxation variable. The separation hyperplane can be obtained by solving α and ω.

The SVM approach handles nonlinear issues by defining an appropriate kernel function, which effectively computes the similarity between samples and landmark points to establish new features to train complicated nonlinear decision limits. There are four kernel functions in SVM: linear, polynomial, hyperbolic tangent (sigmoid), and Gaussian radial basis function (RBF). Because multiple kernel functions were employed to identify hyperplanes across different data dispersion, the kernel function parameters, as well as their accompanying penalty coefficients C and gamma parameters, impact model accuracy.

2.5. Auxiliary Data

Auxiliary data were generated from the DEM and satellite imagery of Sentinel 2A. Terrain attributes such as aspect, elevation, slope, valley depth (VD), cross-sectional curvature (CSC), longitudinal curvature (LC), plane curvature (Plan_C), flow accumulation (Flow Ac), multi-resolution ridge top flatness (MrRTF), multi-resolution valley bottom flatness (MrVBF), and slope length factor (LS-factor) were prepared through SAGA GIS software v. 2.1.4. (System for Automated Geoscientific Analysis) [54] at a 12.5 m resolution [55], and the normalized difference vegetative index (NDVI), soil-adjusted vegetation index (SAVI) [56], enhanced vegetative index (EVI) [57], and brightness index (BI) [58] were also prepared from Sentinel 2A imagery [59], which was acquired on 2 August 2020. After that, all data were resampled to 10×10 m spatial resolution.

2.6. Selection of Important Features

All features obtained from auxiliary data, i.e., terrain attributes and satellite imagery of the sentinel 2A along with soil properties, were processed through a feature selection Boruta

algorithm in the R statistical package. Boruta is a wrapper algorithm built around the RF. The Boruta algorithm formed at least five copies of the original auxiliary dataset along with soil and climatic data, also known as shadow attributes. To remove the correlation between auxiliary data and land suitability classes, the added feature data were shuffled. Then the algorithm was run with RF classifier over the added feature data and computed using the Z score for each variable and corresponding shadow variable. The Z score represents the importance of each variable for land suitability classes. Those variables which score better than the maximum Z score was assigned with a hit and marked "confirmed" and were considered important. Variables with a lower Z score were assigned 0 hits and marked "Rejected" and were considered unimportant [47].

2.7. Machine Learning Algorithms

RF Models in ArcGIS Pro 2.5.0 were built and predictions made using a modification of Leo Breiman's random forest technique, which is a supervised machine learning technique [50]. Predictions were made for categorical variables, as well as continuous variables. Predictor variables can be added to the attribute table of training samples. Raster datasets then determined neighborhood values for usage as auxiliary variables. Predictions were made to either features or a prediction raster, in addition to evaluation of prediction accuracy based on training data. The SVM classifier tool in ArcGIS Pro 2.5.0 was also used as a supervised classification algorithm. For conventional image inputs, the tool in ArcGIS Pro 2.5.0 takes multiband imagery with depth and does SVM classification on a pixel basis, depending on input training feature file, which was utilized in the classification process.

2.8. Model Evaluation

In this study, classification accuracy was assessed for each classified image by using three different classification techniques (WOL, RF, and SVM), which were then computed from two common performance metrics in the confusion matrix (i.e., overall accuracy and kappa index) for each classification. The overall accuracy was evaluated as the ratio between the sum of all correctly classified land suitability classes and all other classes, as shown in Equation (4).

$$Overall\ Accuracy = \frac{\Sigma\ TP + TN}{\Sigma TP + FP + TN + FN} \tag{4}$$

The kappa index was used, which is a robust indicator that incorporates the probability that a class is categorized by chance. It is a simple statistic that quantifies the ratio of all potential occurrences of inclusion or exclusion that are predicted by a model after taking into consideration chance predictions. A higher kappa index, like a higher overall accuracy, represents a strong model performance and is given in Equation (5).

$$Kappa\ index = \frac{1 - Overall\ accuracy}{1 - hypothetical\ probability\ of\ chance\ agreement} \tag{5}$$

3. Results

3.1. Summary of Statistics

Descriptive statistics of soil properties of the study area are shown in Table 1. The mean value of pH was 7.83 and ranged between 6.40 and 8.70, which shows that the average soil of the study area was sodic. The mean value of EC was 0.22 dsm^{-1} and ranged between 0.03 and 0.68 dsm^{-1}, illustrating a low electrical conductivity. The CaCO$_3$ ranged between 0.03 and 0.68 and had a mean of 0.22%, which represents calcareous soil in the study area. The OC ranged between 0.03% and 0.59% which shows that soil fertility was low. The predominant soil textures in the study area were sandy loam, loam, and silty clay loam. The coefficient of variation (CV) of the EC, CaCO$_3$, clay, and sand were high, i.e., greater than 35%, likely owing to variation in topography and parent material. However, the CV for the rest of the soil properties was moderate, i.e., less than 35%. Therefore, point

data used to build a regression (soil physiochemical properties) had a wide range of values in the study area.

3.2. Selected Auxiliary Features

Figures 3 and 4 depict the selection of each auxiliary feature along with soil and climatic data for crop suitability classification of a rainfed olive–maize agroforestry system. Among auxiliary variables, which were used for estimating land suitability classes for olive, elevation had a maximum importance of 70.33%, followed by VD (valley depth, 56.36%), aspect (49.44%), Flow_Ac (flow accumulation; 41.66%), slope (41.68%), MRVBF (multi-resolution valley bottom flatness index; 34.47%), CSC (cross sectional curvature; 29.06%,) MRTF (multi-resolution ridge top flatness index; 32.18%), LS_factor (land slope factor; 33.33%), plan_C (plane curvature; 22.24%), and long_C (longitudinal curvature; 20.31%). Indices like EVI (enhanced vegetative index; 25.03%), SAVI (soil-adjusted vegetative index; 46.67%), NDVI (normalized difference vegetative index; 43.94%), and BI (brightness index; 41.94%) were obtained from satellite imagery of sentinel-2A. Data such as AT (annual temperature; 65.47%) and ARF (annual rainfall; 16.57%) were obtained from climatic data (Table 1). For maize, aspect had a maximum importance of 41.70%, followed by VD (valley depth; 39.21%), elevation at 37.72%, FlowAC (flow accumulation; 33.48%), SAVI (soil-adjusted vegetative index; 32.95%), slope at 35.07%, NDVI (normalized difference vegetative index; 30.88%), MrVBF (multi-resolution valley bottom flatness index; 30.21%), BI (brightness index; 30.11%), soil texture at 28.40%, CSC (cross-sectional curvature; 21.92%), ATGC (annual temperature of the growing cycle; 20.73%), RFGC (annual temperature of the growing cycle; 19.42%), and EVI (enhance vegetative index; 18.99%).

After the data modeling of all these features, they were classified by RF and SVM by dividing the data into 70% sample training dataset, 20% testing dataset, and 10% validation dataset. For the prediction of each suitable classes, auxiliary features like MRTF and MRVBF illustrated steep terrain, where high values represent flat areas [with high land suitability class (S1)], while lower values represent high steep areas [with marginal and non-suitable land class, i.e., (S3) and (N2)]. The greater relative significance of auxiliary data retrieved from the DEM for land suitability class prediction illustrated the influence of topography as one of the major soil formation components of land suitability for rainfed olive and maize. Almost all the mechanisms involved in soil formation were influenced by topography. Many researchers have indicated that topography has a substantial influence on soil attributes. Rezaei et al. [60] observed that slope plays a meaningful part in soil property changes and soil development. Zadeh et al. [61], during estimation of soil organic carbon (SOC), showed that rainfall, valley depth, terrain surface texture, temperature, channel network base level, and terrain vector roughness were the most relevant ancillary factors.

Similarly, Taghizadeh et al. [18] used various auxiliary features for predicting land suitability classes in the Kurdistan province of Iran. The auxiliary features were MrRTF, MrVBF, NDVI, SAVI, BI, valley depth, cross-sectional curvature, plan curvature, and slope, which were used to assess land suitability classes. Other researchers have observed similar findings [38,39]. Nabiollahi et al. [62] used a digital soil mapping method to analyze spatial variation in SOC stocks under land-use change in western Iran and found that the important auxiliary features were: TWI, MRRTF, MRVBF, NDVI, Band 3, and Band 4 of Landsat 8. Suleman et al. [63] conducted land suitability assessments for olive cultivation in Turkey and also used the auxiliary data of elevation to mitigate the effect of lower temperature and snow, which restricts growth period and production of olives. Soil with slope less than 25% is preferred for cultivation, and aspect data were used to identify flat and south-facing areas.
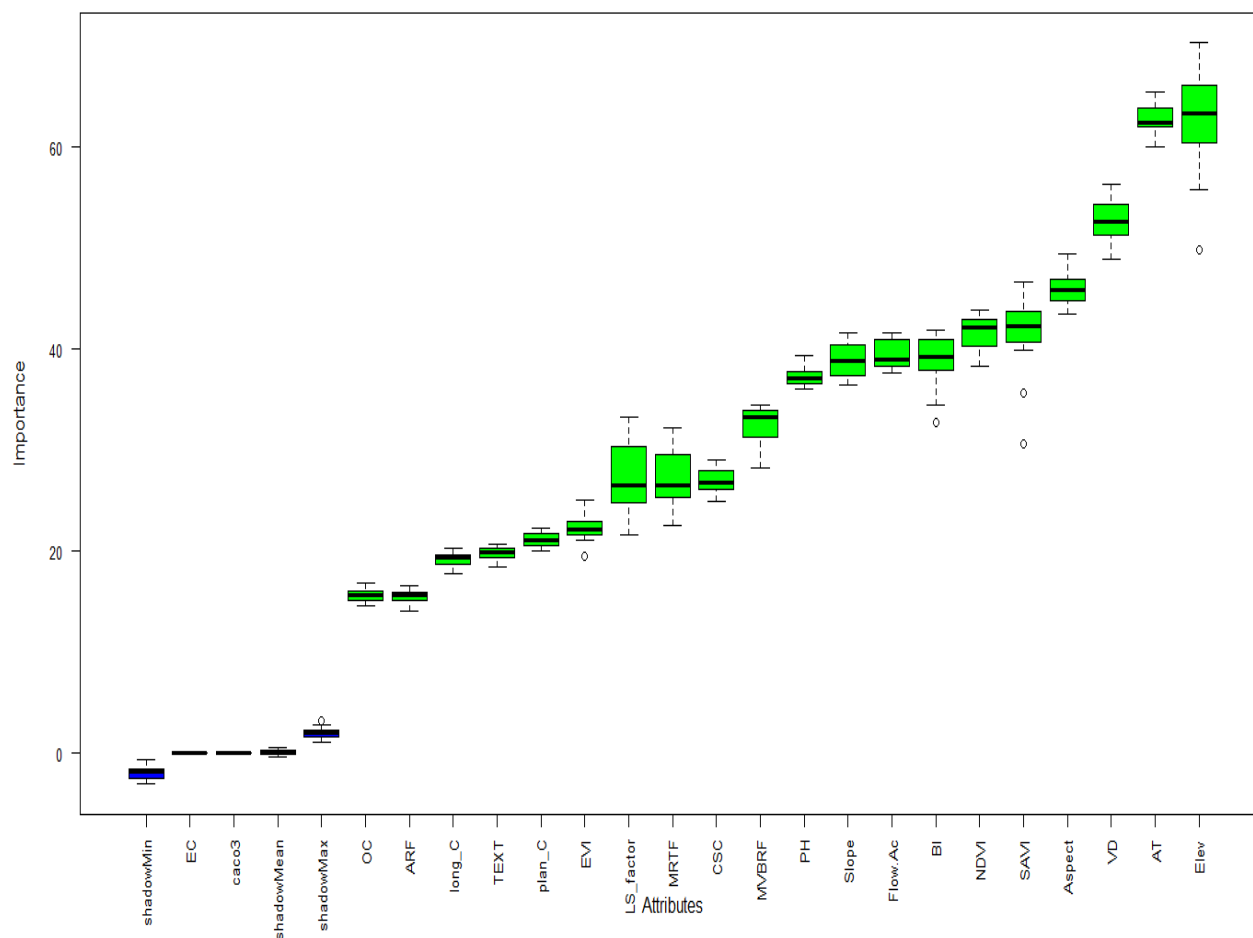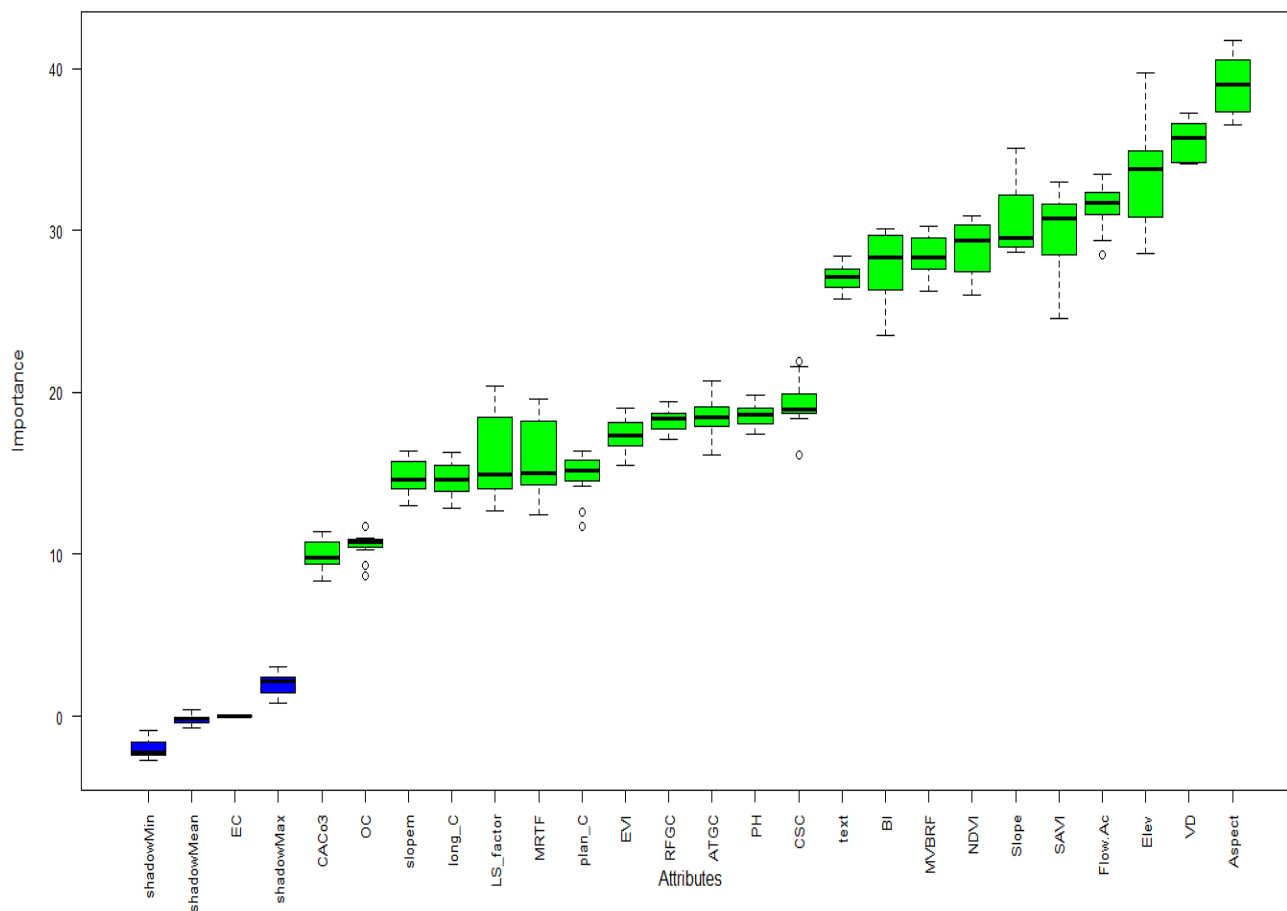
Figure 3. The relative importance of each auxiliary feature along with the soil and climatic data used for prediction of land suitability classes of olive through the Boruta algorithm. (Elev: elevation; AT: annual temperature; VD: valley depth; aspect; SAVI: soil-adjusted vegetative index; NDVI: normalized vegetative index; BI: brightness index; Flow_Ac: flow accumulation; MRTF: multi-resolution ridge top flatness; MVBF: multi-resolution valley bottom flatness; CSC: cross-sectional curvature; LS_factor: land slope factor; EVI: enhanced vegetative index; plan_C: plane curvature; TEXT: Soil texture; long_C: longitudinal curvature; ARF: annual rainfall; OC: organic carbon; shadow min; shadow mean; shadow max; EC: Electrical conductivity; CaCO$_3$: Calcium carbonate).

3.3. Comparison of the ML and Traditional Technique

The performance of two ML models (SVM and RF) and the traditional technique of WOL was employed for the land suitability classification of the agroforestry system of rainfed olive and maize based on overall accuracy and Kappa index (Table 2). The overall accuracy and Kappa index for prediction of land suitability classes for the rainfed olive through RF (0.94 and 0.90, respectively) were higher than SVM (0.92 and 0.88, respectively) and the classification done by WOL (0.91 and 0.87, respectively). On the other hand, the overall accuracy and Kappa index for the land suitability classification for rainfed maize crop performed by RF (0.94 and 0.91, respectively) was higher than SVM (0.93 and 0.90, respectively).

Figure 4. The relative importance of each auxiliary feature along with the soil and climatic data used for prediction of land suitability classes of maize crop through the Boruta algorithm. (Elev: elevation; VD: valley depth; aspect; SAVI: soil-adjusted vegetative index; NDVI: normalized vegetative index; BI: brightness index; Flow_Ac: flow accumulation; MRTF: multi-resolution ridge top flatness; MVBF: multi-resolution valley bottom flatness; CSC: cross-sectional curvature; LS_factor: land slope factor; EVI: enhanced vegetative index; plan_C: plane curvature; TEXT: soil texture; long_C: longitudinal curvature; RFC: rainfall of the growing cycle; ATGC: average temperature of the growing cycle; OC: organic carbon; shadow min; shadow mean; shadow max; EC: Electrical conductivity; $CaCO_3$: Calcium carbonate).

Table 2. Descriptive statistics for soil properties at 30 cm depth and 701 regions.

Soil Properties	Minimum	Maximum	Mean	CV *	SD
pH	6.40	8.70	7.83	6.02	0.47
EC dsm^{-1}	0.03	0.68	0.22	64.21	0.14
$CaCO_3$ %	0.03	0.68	0.22	64.22	0.14
OC %	0.32	0.59	0.45	12.61	0.06
Clay %	2.00	42.00	18.48	41.87	7.74
Silt %	2.00	64.00	41.89	23.36	9.79
Sand %	6.00	84.00	39.63	38.70	15.34

* The coefficient of variation (CV) is the dispersion of the data points around the average. The standard deviation (SD) must always be considered in the context of the mean of the data.

Based on overall accuracy and the Kappa Index [64,65], the two ML algorithms (RF and SVM) and the traditional technique of WOL had strong, moderate, and poor capabilities for predicting soil and land suitability classes for a rainfed olive–maize agroforestry system (Table 3). Statistically, the performance of RF was better than the other two methods in terms of predicting suitability classes. When compared to single classifiers, ensemble

learning techniques (e.g., RF) indicated effectiveness and robustness to noise and could handle a large amount of input data [66].

Table 3. Error criteria used for the land suitability classification for agroforestry system of olive and maize crop through ML (RF: random forest; SVM: support vector machine) and traditional technique (WOL: weighted overlay).

ML and Traditional Techniques	Overall Accuracy		Kappa Index	
	Rainfed Olive	Rainfed Maize	Rainfed Maize	Rainfed Olive
RF	0.94	0.94	0.91	0.90
SVM	0.92	0.93	0.90	0.88
WOL	0.91	0.87	0.85	0.87

3.4. Land Suitability Class

3.4.1. Highly Suitable Class (S1)

The area classified as highly suitable through ML (RF and SVM) and traditional method (WOL) was 5%, 9%, and 25%, respectively, of the total area for olive and about 6%, 7%, and 21%, respectively, of the total area classified in the highly suitable class (S1). The area of S1 class had ideal conditions for an olive and maize agroforestry system. This S1 categorization was due to lower slope, optimal temperature and rainfall amount, highly suitable soil pH, and appropriate soil texture, while only the SOC was less than the required range but could be augmented by adding manure or fertilizers. Most of the S1 class mainly represents valley areas, which have almost no limitation for such a system and do not require any special treatment for soil improvement. A study conducted by Mukhtar Elaalem [67], through the comparison of parametric and fuzzy multicriteria methods for evaluating land suitability for olive in Libya, classified the area of S1 as 30% (92,819 ha), S2 as 50% (154,698 ha), S3 as 6% (18,564 ha), and N1 which was 9% (27,846 ha).

3.4.2. Moderately Suitable Class (S2)

The rating scale for the S2 class between ML algorithms ranged between 65 and 85 [48]. The percentage of area which was classified as S2 through RF, SVM, and WOL for olive production was 21%, 21%, and 57%, respectively, while the area classified for the maize crop through RF, SVM, and WOL was 34%, 49%, and 63%, respectively, of the total area. As compared to the S1 class, the S2 class had slight limitations owing to slope, elevation, pH, SOC, temperature, and rainfall. Soil texture was optimal for olive but had limitation for maize and thus was classified into S2. The majority of these limitations can be overcome by management; for example, slope could be improved with the construction of terraces, SOC could be improved by adding manure or green manure, and pH could be reduced by adding lime in S2 areas [68].

3.4.3. Marginally Suitable Class (S3)

The rating scale of the S3 class ranged between 40 and 60 [48]. The percentage of the area which was classified under this class through RF, SVM, and WOL for olive was 56%, 51%, and 16%, respectively, while the percentage of the area classified through RF, SVM, and WOL for maize was 40%, 27%, and 16%, respectively. The limitations of soil physicochemical attributes, topography features, and climatic factors were lower than in the S2 class, and thus the potential success of olive and maize crops was lower than in the S2 class. Since, in the S3 class, slopes are steep and the area is largely located on high elevation where widespread watering is not viable during times of drought, these areas would not be optimum for integrated maize–olive systems (Table 4).

Table 4. Area of the land suitability classes in hectare for the agroforestry system of olive and maize crop through ML (RF, SVM) and traditional technique of weighted overlay.

| Land Suitability Class | ML-Based Technique | | | | | | | | Traditional Technique | | | |
| | RF Olive | | SVM Olive | | RF Maize | | SVM Maize | | WOL Olive | | WOL Maize | |
	Area (ha)	Area (%)	Area (ha)	Area (%)	Area (ha)	Area (%)	Area (ha)	Area (%)	Area (ha)	Area (%)	Area (ha)	Area (%)
Highly suitable (S1)	9567	5.00	14,924	9.00	10,673	6.00	11,543	7.00	36,167	25.00	43,756	21.00
Moderately suitable (S2)	36,171	21.00	36,131	21.00	60,336	34.00	85,835	49.00	110,835	57.00	99,480	63.00
Marginally suitable (S3)	97,488	56.00	90,146	51.00	70,366	40.00	48,525	27.00	28,134	16.00	29,171	16.00
Permanently non-suitable (N2)	32,165	18.00	34,165	19.00	34,006	20.00	29,480	17.00	3002	2.00	56	0.00

3.4.4. Permanently Non-Suitable Class (N2)

The rating scale for N2 ranged between 0 and 30 [48]. The percentage of area which was classified as N2 through RF, SVM, and WOL for olive was 18%, 19%, and 2%, respectively, while the area classified through RF, SVM, and WOL for maize was 20%, 17%, and 0%, respectively. This area has the limitations of steepness, shallow soil depth, high pH, and low temperature. Such areas cannot be improved by any means and should be avoided for plantation cultivation. The overall land suitability classification for the olive and maize agroforestry system in rainfed areas is shown in Figure 5. Ostovari et al. [69] used GIS and multi-criteria decision-making for analyzing and creating a land suitability map of 51,831 ha in the northwest region of East Azerbaijan province, Iran for rapeseed (*Brassica napus*). Their findings revealed that 0.81% (420.8 ha), 42.3% (21,940 ha), and 11.7% (6104 ha) of the study region were highly suitable (S1), moderately suitable (S2), and marginally suitable (S3), respectively, and 39.72% (20,586) and 0.95% (492.1 ha) were currently and permanently unsuitable for rapeseed cultivation. In their studied region, soil characteristics and topographical variables (elevation and slope) were the most influential variables for identifying the N2 class.
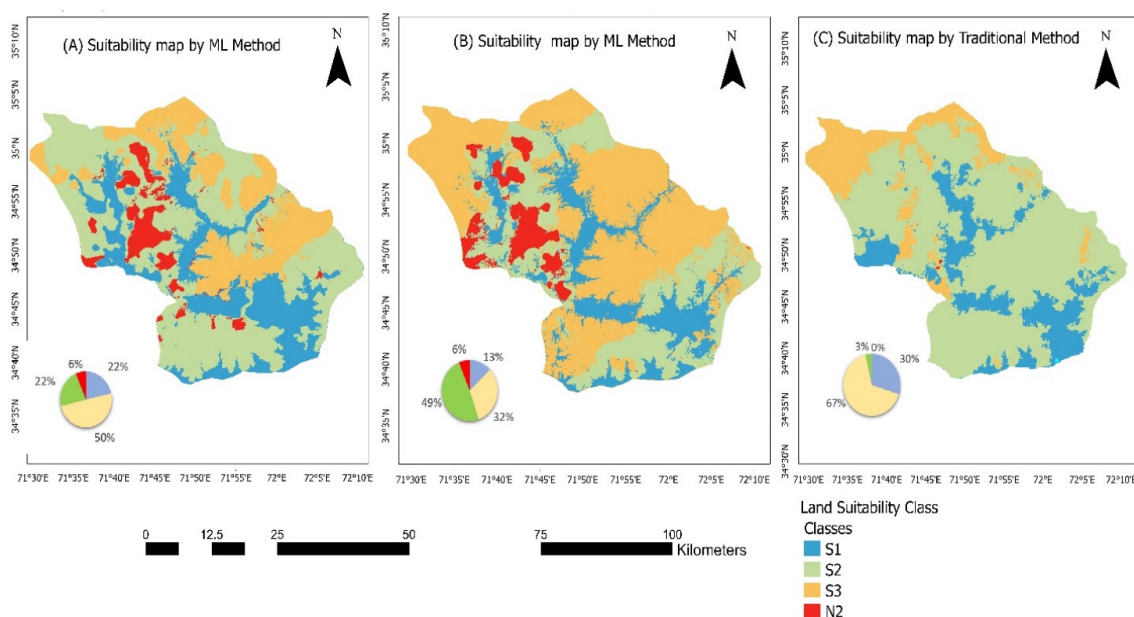
Figure 5. Land suitability maps for an integrated olive–maize agroforestry system in rainfed areas through ML algorithm (**A**) random forest (RF), (**B**) support vector machine (SVM), and through traditional method of (**C**) weighted overlay (WOL).

4. Discussion

Land suitability classification for an integrated olive–maize agroforestry system predicted four classes in the study area: highly suitable area (S1), moderately suitable area (S2), marginally suitable area (S3), and permanently non-suitable area (N2). The classification was performed by ML methods (RF and SVM) and a traditional method (WOL). Statically, the ML RF method performed better than the SVM and weighted overlay WOL. The area of S1 was predicted through RF to be 9%↑ than that of SVM and was 8%↓ than the area predicted by WOL. The area of S2 predicted through RF was 18%↑ than that of SWM and was 17%↓ than the area predicted by WOL. Similarly, the area of S3 class predicted by RF was 27%↓ than that of SVM, and 45%↓ than the area predicted by WOL. On the other hand, the area of N2 predicted by RF and SVM was 6%↑ than area predicted through WOL. Thus, the WOL technique under-predicted the N2 class and over-predicted the S2 class. Similar results were produced by Mubashir et al. [6] during their land suitability analysis for sugar cane (*Saccharum officinarum*) in the Bijnor district in India.

From this research, it was confirmed that a sizeable area of the district had great potential for an integrated olive and maize agroforestry system, and that the unrealized area (currently uncultivated) could be used for cultivating this system. Therefore, by using the maps of this study, farmers, especially those who are involved in maize farming, can also incorporate olive trees within rows to enhance profitability and ecosystem services. Such a plantation of olive can also fulfill fodder needs and reduce the pressure on existing forest resources for fuelwood requirements. Thus, such dual-use cropping systems will enhance land productivity and help meet demands of the local population by ensuring food security in this mountainous region.

Maps produced by ML are recommended for land suitability evaluation. This is especially true in a country like Pakistan where soil data are scarce. As a result, ML and ancillary data may be an appealing strategy for large-scale land suitability analyses. From the results, maps produced through ML algorithms predicted more area of N2 and S3 classes compared with weighted overlay. However, the weighted overlay predicted more area of S1 and S2 classes than the ML technique. Future studies could evaluate the potential production of an integrated olive–maize agroforestry system. Finally, the same method presented in this paper could be conducted for land suitability assessments of different crops.

Climate change and shifting agricultural land use threaten agriculture production and food security. Agroforestry systems can enhance land productivity while reducing climate change effects, such as through improved soil moisture loss and reduced below-canopy temperature [38,39]. Production potential from rainfed agriculture is low, but this dual-use cropping system may increase productivity and improve the livelihood of farmers. Future land suitability assessments can be used to analyze suitable areas for various agroforestry systems in different regions of Pakistan according to the suitability of climate, soil, and topography. Future additional research may include implementing the establishment of maize–olive in highly suitable land and verifying responses, as well as conducting additional agroforestry intercropping suitability assessments for the region under investigation.

5. Conclusions

This research identified suitable land for an olive–maize intercropped agroforestry system in rainfed areas for diversifying systems and markets. The elevation, valley depth, aspect, SAVI, NDVI, BI, flow accumulation, MRTF, MVBF, cross-sectional curvature, land slope factor, EVI, plane curvature, and longitudinal curvature were the most important ancillary data. Considering the Kappa index and overall accuracy scores, the RF algorithm outperformed SVM and WOL in determining land suitability of an integrated olive–maize rainfed agroforestry system in the Khyber Pakhtunkhwa district. Overall, RF provided several advantages over other modeling techniques and was deemed optimum for determining land suitability classes.

Results indicate that the region under investigation was largely suitable for an intercropping maize–olive agroforestry system due to optimal rainfall and temperature during the vegetative growth stage, moderate uphill slopes, adequate soil depth, and a suitable pH. Land improvement activities such as terracing, increasing soil organic matter, supplementary irrigation, and lime would be required to improve the suitability of moderately and marginally suitable areas and increase their potential production. This study offers valuable information that may be useful to policymakers and land managers for increasing valuable olive–maize systems in the Khyber Pakhtunkhwa province and adjacent areas.

References

1. Hsu, H.-H.; Chern, W.S.; Gale, F. How will rising income affect the structure of food demand? In *China's Food and Agriculture: Issues for the 21st Century*; USDA: Washington, DC, USA, 2001; pp. 10–13.
2. Xiong, B.; Sumner, D.; Matthews, W. A new market for an old food: The US demand for olive oil. *Agric. Econ.* **2014**, *45*, 107–118. [CrossRef]
3. Zaidi, S.M.H. Edible oil imports in Pakistan. *S. Asian J. Manag.* **2014**, *8*, 1–8.
4. Lutz, W.; Sanderson, W.; Scherbov, S. The end of world population growth. *Nature* **2001**, *412*, 543–545. [CrossRef]
5. Rahman, S.A.; Jacobsen, J.B.; Healey, J.R.; Roshetko, J.M.; Sunderland, T. Finding alternatives to swidden agriculture: Does agroforestry improve livelihood options and reduce pressure on existing forest? *Agrofor. Syst.* **2017**, *91*, 185–199. [CrossRef]
6. Jamil, M.; Ahmed, R.; Sajjad, H. Land suitability assessment for sugarcane cultivation in Bijnor district, India using geographic information system and fuzzy analytical hierarchy process. *GeoJournal* **2018**, *83*, 595–611. [CrossRef]
7. Tariq, M.; Iqbal, H. Maize in Pakistan—An overview. *Agric. Nat. Resour.* **2010**, *44*, 757–763.
8. Viswanath, S.; Sandhya, M. Traditional agroforestry systems and practices: A review. *Adv. Agric. Res. Technol. J.* **2018**, *2*, 18–29.
9. Jiang, X.J.; Liu, W.; Wu, J.; Wang, P.; Liu, C.; Yuan, Z.Q. Land Degradation Controlled and Mitigated by Rubber-based Agroforestry Systems through Optimizing Soil Physical Conditions and Water Supply Mechanisms: A Case Study in Xishuangbanna, China. *Land Degrad. Dev.* **2017**, *28*, 2277–2289. [CrossRef]
10. Zou, X.; Sanford, R.L. Agroforestry systems in China: A survey and classification. *Agrofor. Syst.* **1990**, *11*, 85–94. [CrossRef]
11. Dagar, J. Agroforestry systems for the Andaman and Nicobar Islands. *Int. Tree Crops J.* **1995**, *8*, 107–128. [CrossRef]
12. May, P.H.; Anderson, A.B.; Frazão, J.M.F.; Balick, M.J. Babassu palm in the agroforestry systems in Brazil's Mid-North region. *Agrofor. Syst.* **1985**, *3*, 275–295. [CrossRef]
13. Paris, P.; Camilli, F.; Rosati, A.; Mantino, A.; Mezzalira, G.; Dalla Valle, C.; Franca, A.; Seddaiu, G.; Pisanelli, A.; Lauteri, M.; et al. What is the future for agroforestry in Italy? *Agrofor. Syst.* **2019**, *93*, 2243–2256. [CrossRef]
14. Hayat, M.; Zha, T.; Nizami, S.M.; Gulzar, S.; Khan, A.; Iqbal, S.; Khan, M.S. Productive role of agroforestry system in context of ecosystem services in district dir lower. pakistan. *Pak. J. Bot.* **2020**, *52*, 1411–1419. [CrossRef] [PubMed]
15. FAO. *Statistical Yearbook*; FAO: Rome, Italy, 2013.
16. Rossiter, D.G. Land evaluation: Towards a revised framework; Land and Water Discussion Paper 6, FAO. FAO, Rome (2007). *Geoderma* **2009**, *148*, 428–429. [CrossRef]
17. Halder, J.C. Land suitability assessment for crop cultivation by using remote sensing and GIS. *J. Geogr. Geol.* **2013**, *5*, 65. [CrossRef]
18. Taghizadeh-Mehrjardi, R.; Nabiollahi, K.; Rasoli, L.; Kerry, R.; Scholten, T. Land suitability assessment and agricultural production sustainability using machine learning models. *Agronomy* **2020**, *10*, 573. [CrossRef]
19. FAO. *A Framework for Land Evaluation*; FAO and Agriculture Organization of the United Nations: Rome, Italy, 1976.
20. Sys, C.; Van Ranst, E.; Debaveye, J. Land Evaluation. Part III: Crop Requirements. Agricultural Publications N° 7, G.A.D.C., Brussels, Belgium. 1993. Available online: https://biblio.ugent.be/publication/233235 (accessed on 2 December 2020).
21. Mustafa, A.; Singh, M.; Sahoo, R.; Ahmed, N.; Khanna, M.; Sarangi, A.; Mishra, A. Land suitability analysis for different crops: A multi criteria decision making approach using remote sensing and GIS. *Researcher* **2011**, *3*, 61–84.
22. Hassan, I.; Javed, M.A.; Asif, M.; Luqman, M.; Ahmad, S.R.; Ahmad, A.; Akhtar, S.; Hussain, B. Weighted overlay based land suitability analysis of agriculture land in Azad Jammu and Kashmir using GIS and AHP. *Pak. J. Agric. Sci.* **2020**, *57*, 1509–1519.

23. Perveen, S.; Arsalan, M.H.; Siddiqui, M.F.; Khan, I.A.; Anjum, S.; Abid, M. GIS-based multi-criteria model for cotton crop land suitability: A perspective from Sindh province of Pakistan. *FUUAST J. Biol.* **2013**, *3*, 31–37.
24. Naz, A.; Rasheed, H. Modeling the rice land suitability using GIS and multi-criteria decision analysis approach in Sindh, Pakistan. *J. Basic Appl. Sci.* **2017**, *13*, 26–33.
25. Behrens, T.; Scholten, T. Digital soil mapping in Germany—A review. *J. Plant Nutr. Soil Sci.* **2006**, *169*, 434–443. [CrossRef]
26. Hengl, T.; Mendes de Jesus, J.; Heuvelink, G.B.; Ruiperez Gonzalez, M.; Kilibarda, M.; Blagotić, A.; Shangguan, W.; Wright, M.N.; Geng, X.; Bauer-Marschallinger, B. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE* **2017**, *12*, e0169748. [CrossRef] [PubMed]
27. Mokarram, M.; Hamzeh, S.; Aminzadeh, F.; Zarei, A.R. Using machine learning for land suitability classification. *West Afr. J. Appl. Ecol.* **2015**, *23*, 63–73.
28. Ghiasi, R.; Ghasemi, M.R.; Noori, M. Comparative studies of metamodeling and AI-Based techniques in damage detection of structures. *Adv. Eng. Softw.* **2018**, *125*, 101–112. [CrossRef]
29. Fidêncio, P.H.; Poppi, R.J.; de Andrade, J.C.; Cantarella, H. Determination of organic matter in soil using near-infrared spectroscopy and partial least squares regression. *Commun. Soil Sci. Plant Anal.* **2002**, *33*, 1607–1615. [CrossRef]
30. Gruba, P.; Socha, J. Exploring the effects of dominant forest tree species, soil texture, altitude, and pHH2O on soil carbon stocks using generalized additive models. *For. Ecol. Manag.* **2019**, *447*, 105–114. [CrossRef]
31. Puente, C.; Olague, G.; Trabucchi, M.; Arjona-Villicaña, P.D.; Souberville-Montalvo, C. Synthesis of vegetation indices using genetic programming for soil erosion estimation. *Remote Sens.* **2019**, *11*, 156. [CrossRef]
32. Bui, D.T.; Bui, Q.-T.; Nguyen, Q.-P.; Pradhan, B.; Nampak, H.; Trinh, P.T. A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area. *Agric. For. Meteorol.* **2017**, *233*, 32–44.
33. Azeez, N.; Yahya, W.; Al-Taie, I.; Basbrain, A.; Clark, A. Regional Agricultural Land Classification Based on Random Forest (RF), Decision Tree, and SVMs Techniques. In *Fourth International Congress on Information and Communication Technology*; Springer: Singapore, 2020; pp. 73–81.
34. Löw, F.; Schorcht, G.; Michel, U.; Dech, S.; Conrad, C. Per-field crop classification in irrigated agricultural regions in middle Asia using random forest and support vector machine ensemble. In *Earth Resources and Environmental Remote Sensing/GIS Applications III*; SPIE: Bellingham, WA, USA, 2012; p. 85380R.
35. McBratney, A.B.; Santos, M.M.; Minasny, B. On digital soil mapping. *Geoderma* **2003**, *117*, 3–52. [CrossRef]
36. Ranjitkar, S.; Sujakhu, N.M.; Lu, Y.; Wang, Q.; Wang, M.; He, J.; Mortimer, P.E.; Xu, J.; Kindt, R.; Zomer, R.J. Climate modelling for agroforestry species selection in Yunnan Province, China. *Environ. Model. Softw.* **2016**, *75*, 263–272. [CrossRef]
37. Chen, Y.; Wu, B.; Chen, D.; Qi, Y. Using machine learning to assess site suitability for afforestation with particular species. *Forests* **2019**, *10*, 739. [CrossRef]
38. Jiang, Z.-D.; Owens, P.R.; Ashworth, A.J.; Fuentes, B.A.; Thomas, A.L.; Sauer, T.J.; Wang, Q.-B. Evaluating tree growth factors into species-specific functional soil maps for improved agroforestry system efficiency. *Agrofor. Syst.* **2021**, *96*, 479–490. [CrossRef]
39. Kharel, T.P.; Ashworth, A.J.; Owens, P.R.; Philipp, D.; Thomas, A.L.; Sauer, T.J. Teasing apart silvopasture system components using machine learning for optimization. *Soil Syst.* **2021**, *5*, 41. [CrossRef]
40. Tan, K.H. *Soil Sampling, Preparation, and Analysis*; CRC Press: Boca Raton, FL, USA, 1995.
41. Carter, M.R.; Gregorich, E.G. *Soil Sampling and Methods of Analysis*; CRC Press: Boca Raton, FL, USA, 2007.
42. Beretta, A.N.; Silbermann, A.V.; Paladino, L.; Torres, D.; Bassahun, D.; Musselli, R.; García-Lamohte, A. Soil texture analyses using a hydrometer: Modification of the Bouyoucos method. *Int. J. Agric. Nat. Resour.* **2014**, *41*, 263–271. [CrossRef]
43. Allison, L.E. Wet-combustion apparatus and procedure for organic and inorganic carbon in soil. *Soil Sci. Soc. Am. J.* **1960**, *24*, 36–40. [CrossRef]
44. Zhang, R.; Wienhold, B.J. The effect of soil moisture on mineral nitrogen, soil electrical conductivity, and pH. *Nutr. Cycl. Agroecosystems* **2002**, *63*, 251–254. [CrossRef]
45. Chaney, R.; Slonim, S.; Slonim, S. Determination of calcium carbonate content in soils. In *Geotechnical Properties, Behavior, and Performance of Calcareous Soils*; ASTM International: West Conshohocken, PA, USA, 1982.
46. Eckert, D.; Sims, J.T. Recommended soil pH and lime requirement tests. *Recomm. Soil Test. Proced. Northeast. USA Northeast Reg. Bull.* **1995**, *493*, 11–16.
47. Kursa, M.B.; Rudnicki, W.R. Feature selection with the Boruta package. *J. Stat. Softw.* **2010**, *36*, 1–13. [CrossRef]
48. Sys, C.; Van Ranst, E.; Debavey, J.; Beernaert, F. *Land Evaluation, Part III Crop Requirements Agricultural Publication No. 7 General Administration for Development Cooperation*; Ghent University: Ghent, Belgium, 1993.
49. Mbwambo, S.; Maro, G.; Monyo, H.; Mosi, E. Towards Expansion of Coffea canephora Production in Tanzania: The Land Suitability Perspective. *World* **2020**, *8*, 52–56.
50. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
51. Tiwari, A.; Ajmera, S. Land Suitability Assessment for Agriculture Using Analytical Hierarchy Process and Weighted Overlay Analysis in ArcGIS ModelBuilder. In *Recent Trends in Civil Engineering*; Springer: Singapore, 2021; pp. 735–762.
52. Yalew, S.G.; Van Griensven, A.; Mul, M.L.; van der Zaag, P. Land suitability analysis for agriculture in the Abbay basin using remote sensing, GIS and AHP techniques. *Model. Earth Syst. Environ.* **2016**, *2*, 101. [CrossRef]

53. Shaheen, A.; Iqbal, J. Spatial distribution and mobility assessment of carcinogenic heavy metals in soil profiles using geostatistics and random forest, boruta algorithm. *Sustainability* **2018**, *10*, 799. [CrossRef]
54. Olaya, V. *A Gentle Introduction to SAGA GIS*; The SAGA User Group eV: Gottingen, Germany, 2004; Volume 208.
55. *Dataset: ASF DAAC 2015, ALOS PALSAR_Radiometric_Terrain_Corrected_High_Res*; ASF: Fairbanks, AK, USA, 2011. [CrossRef]
56. Habibie, M.I.; Noguchi, R.; Shusuke, M.; Ahamed, T. Land suitability analysis for maize production in Indonesia using satellite remote sensing and GIS-based multicriteria decision support system. *GeoJournal* **2019**, *86*, 777–807. [CrossRef]
57. Hossain, M.S.; Chowdhury, S.R.; Das, N.G.; Rahaman, M.M. Multi-criteria evaluation approach to GIS-based land-suitability classification for tilapia farming in Bangladesh. *Aquac. Int.* **2007**, *15*, 425–443. [CrossRef]
58. Apriliani, A.; Kusumaningrum, R.; Endah, S.N.; Prasetyo, Y. Suitability analysis of rice varieties using learning vector quantization and remote sensing images. *Telkomnika* **2019**, *17*, 1290–1299. [CrossRef]
59. Sentinel-2A (ESA) Image Courtesy of the U.S. Geological Survey. 2020. Available online: https://www.usgs.gov/centers/eros/science/usgs-eros-archive-sentinel-2?qt-science_center_objects=0#qt-science_center_objects (accessed on 15 January 2021).
60. Rezaei, H.; Jafarzadeh, A.A.; Alijanpour, A.; Shahbazi, F.; Kamran, K.V. Effect of slope position on soil properties and types along an elevation gradient of Arasbaran forest, Iran. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2015**, *5*, 449–456. [CrossRef]
61. Mahmoudzadeh, H.; Matinfar, H.R.; Taghizadeh-Mehrjardi, R.; Kerry, R. Spatial prediction of soil organic carbon using machine learning techniques in western Iran. *Geoderma Reg.* **2020**, *21*, e00260. [CrossRef]
62. Nabiollahi, K.; Eskandari, S.; Taghizadeh-Mehrjardi, R.; Kerry, R.; Triantafilis, J. Assessing soil organic carbon stocks under land-use change scenarios using random forest models. *Carbon Manag.* **2019**, *10*, 63–77. [CrossRef]
63. Bilgilioğlu, S.S. Land suitability assessment for Olive cultivation using GIS and multi-criteria decision-making in Mersin City, Turkey. *Arab. J. Geosci.* **2021**, *14*, 2434. [CrossRef]
64. Gao, Y.; Zhang, A.; Yue, Y.; Wang, J.a.; Su, P. Predicting Shifts in Land Suitability for Maize Cultivation Worldwide Due to Climate Change: A Modeling Approach. *Land* **2021**, *10*, 295. [CrossRef]
65. Heumann, B.W.; Walsh, S.J.; McDaniel, P.M. Assessing the application of a geographic presence-only model for land suitability mapping. *Ecol. Inform.* **2011**, *6*, 257–269. [CrossRef] [PubMed]
66. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
67. Elaalem, M. A comparison of parametric and fuzzy multi-criteria methods for evaluating land suitability for olive in Jeffara Plain of Libya. *Apcbee Procedia* **2013**, *5*, 405–409. [CrossRef]
68. McCauley, A.; Jones, C.; Jacobsen, J. *Soil pH and Organic Matter, Nutrient Management'S Module*; Montana State University: Bozeman, MT, USA, 2009.
69. Kazemi, H.; Akinci, H. A land use suitability model for rainfed farming by Multi-criteria Decision-making Analysis (MCDA) and Geographic Information System (GIS). *Ecol. Eng.* **2018**, *116*, 1–6. [CrossRef]