



# Article Multi-Temporal Sentinel-1 and Sentinel-2 Data for Orchards Discrimination in Khairpur District, Pakistan Using Spectral Separability Analysis and Machine Learning Classification

Arif Ur Rehman<sup>1,2</sup>, Lifu Zhang<sup>1,2,\*</sup>, Meer Muhammad Sajjad<sup>2,3</sup> and Abdur Raziq<sup>4</sup>

- <sup>1</sup> Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; arifrehman07@mails.ucas.ac.cn
- <sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, China; meersajjad@igsnrr.ac.cn
- <sup>3</sup> State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
- <sup>4</sup> Department of Geography, Islamia College University, Peshawar 25120, Pakistan; drraziq@icp.edu.pk
- \* Correspondence: zhanglf@aircas.ac.cn

**Abstract:** Generating orchards spatial distribution maps within a heterogeneous landscape is challenging and requires fine spatial and temporal resolution images. This study examines the effectiveness of Sentinel-1 (S1) and Sentinel-2 (S2) satellite data of relatively high spatial and temporal resolutions for discriminating major orchards in the Khairpur district of the Sindh province, Pakistan using machine learning methods such as random forest (RF) and a support vector machine. A Multicollinearity test (MCT) was performed among the multi-temporal S1 and S2 variables to remove those with high correlations. Six different feature combination schemes were tested, with the fusion of multi-temporal S1 and S2 (scheme-6) outperforming all other combination schemes. The spectral separability between orchards pairs was assessed using Jeffries-Matusita (JM) distance, revealing that orchard pairs were completely separable in the multi-temporal fusion of both sensors, especially the indistinguishable pair of dates-mango. The performance difference between RF and SVM was not significant, SVM showed a slightly higher accuracy, except for scheme-4 where RF performed better. This study concludes that multi-temporal fusion of S1 and S2 data, coupled with robust ML methods, offers a reliable approach for orchard classification. Prospectively, these findings will be helpful for orchard monitoring, improvement of yield estimation and precision based agricultural practices.

**Keywords:** multi-temporal; spectral separability; orchard classification; random forest; support vector machine

# 1. Introduction

Fruit orchards play a vital role in global and regional economies, serving as cash crops with direct economic impacts. Accurate mapping of fruit orchards is crucial for research on their disease prevention, pest control, yield estimation, sustainable agricultural management and other socio-economic concerns [1]. Traditionally, orchard mapping requires time-consuming, labor-intensive and costly field surveys [2,3]. However, the development of remote sensing technologies allows for efficient mapping at low cost and at different geographic scales [4,5]. Remote sensing images, taking into account the spectral and temporal characteristics, are widely used to discriminate vegetation types based on their spectral signatures at different time periods [6,7]. Multi-temporal remote sensing images provide crucial time-based spectral information, particularly significant for orchard classification.

Compared to Landsat and MODIS satellites, the Sentinel-1 (S1) satellite, Synthetic Aperture Radar (SAR) of its C-band dataset and Sentinel-2 (S2) optical satellites have the advantage of their high spatial, temporal and spectral resolution, offering unique



Citation: Rehman, A.U.; Zhang, L.; Sajjad, M.M.; Raziq, A. Multi-Temporal Sentinel-1 and Sentinel-2 Data for Orchards Discrimination in Khairpur District, Pakistan Using Spectral Separability Analysis and Machine Learning Classification. *Remote Sens.* **2024**, *16*, 686. https://doi.org/10.3390/ rs16040686

Academic Editors: Liangliang Tao, Dongryeol Ryu and Hao Sun

Received: 28 December 2023 Revised: 9 February 2024 Accepted: 12 February 2024 Published: 15 February 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). opportunities for global vegetation mapping [8]. Fusion of multi-temporal and multi sensor (MTMS) remote sensing images provides more appropriate datasets for vegetation types mapping [9,10]. Considering spectral similarity between vegetation types, single-date imagery may be unreliable [11,12]. Numerous studies have attempted to classify tree species using MTMS images [13–16], especially combining optical and SAR datasets [17,18]. The effective combination of optical and SAR datasets has shown advantages in vegetation classification [19]. Pham et al. [20] tested S1 alone, fusion of S1 and S2 datasets and only S2 to assess the land cover (LC) classification in a complicated landscape of agricultural in the Vietnam coastal area utilizing RF and OBIA algorithms. The findings indicated that using S1 alone had a lower accuracy result, with an accuracy rate of 60% against 63% and 72% accuracy rate, for S2 alone, as well as fused S1 and S2, respectively. The same conclusion was confirmed by [21]; both individual S1 and S2 produce lower accuracies as compared to the fusion of S1, S2 and Planet-Scope data, producing higher accuracy with a kappa coefficient greater than 0.8%. Similarly, Steinhausen et al. [22], gathered the optical and SAR data for high classification outputs. The utilization of optical and SAR datasets for vegetation mapping depends on the EM spectrum which is highly sensitive to the leaf structure, water and chlorophyll contents. SAR backscattering affects the roughness, structure, texture and the plant canopy dielectric properties [23]. Thus, combining these two datasets provides complementary information and improves accuracy in LC classification.

Accurate LC classification relies on advanced machine learning (ML) methods [24]. Thus, the selection of the finest classification method still remains a topic of ongoing debate within the remote sensing community [25]. Supervised ML methods, like Random Forest (RF) and Support Vector Machine (SVM) have shown high reliability and accuracy in remote sensing applications, especially in tree species classification tasks [26–30]. However, the performance variation among ML methods is often attributed to different combinations of remote sensing variables. While MTMS datasets enhance classification, they can be high-dimensional, redundant and multicollinear, leading to an increase in computational time and require more storage [31]. The inclusion of highly correlated variables can make the model sensitive to multicollinearity which affect the performance of the classification model, as well as reduce the precision and accuracy. Selecting the most useful variables from MTMS datasets yields more accurate results than using all the available variables. This was supported by [32,33], showing that images acquired in peak growth seasons are more correlated than images acquired in low growth seasons. Variable selection techniques, such as multicollinearity tests (MCT), help optimize the selection of variables for classification by removing highly correlated variables [31]. A threshold value of 0.7 is often considered appropriate for variable removal [34,35].

Pakistan is a rich country in terms of food production, particularly in major orchards contributing significantly to global exports. Notably, bananas, dates and mangoes play a pivotal role. Bananas hold a key position in the agriculture economy, covering an area of 34,800 hectares with a production of 154,800 tons [36]. Sindh province, where the soil and climatic conditions are favorable for banana cultivation, contributes 82%, while Punjab, Khyber Pakhtunkhwa and Baluchistan share 5%, 9.9% and 3%, respectively [37]. Mango, ranking high in both area and production, covers 156,000 hectares and 1,753,000 tons. Pakistan stands as the fourth largest mango producing and exporting country in the world [38]. Mango cultivation is prominent in Punjab and Sindh, recognized for the best quality mangoes in the world [39–41]. In dates production, Pakistan is the seventh largest producer with 557,279 metric tons in 2011, also leading in exports [42,43]. Dates are cultivated in all four provinces of Pakistan, with major producing districts including Khairpur and Sukkar (Sindh), Panjgur (Balochistan), Jhang, Muzaffar Garh and Multan (Punjab) and Dera Ismail Khan (Khyber Pakhtunkhwa). Sindh, particularly the Khairpur district, contributes significantly, accounting for approximately 85% of the total production with an area of 24,425 hectares yielding 164,653 tons [44,45].

Generally, several studies have been conducted using MTMS datasets for tree species classification. However, in Pakistan, no study has been conducted on orchards clas-

sification using MTMS images. Therefore, this study aims to utilize fine resolution multi-temporal S1 SAR and S2 optical images for orchard classification, focusing on the Khairpur district of Pakistan. Two well-known ML algorithms, RF and SVM, were applied to classify major orchard types including bananas, dates and mangoes. This study provides a relevant basis for applying fine spatial resolution, multi-temporal SAR and optical datasets for orchard classification, as well as supporting planting subsidies and production layout planning. The key objectives include (1) orchard discrimination with single and multi-temporal SAR and optical datasets using the JM Distance test, (2) comparative assessment of orchard classification abilities using RF and SVM methods, (3) evaluation of the significance of combining multi-temporal SAR and optical variables, along with vegetation indices, to improve orchard classification accuracy. The research outcomes are expected to provide a foundation for future applications of multi-temporal SAR and optical datasets in orchard classification.

#### 2. Materials and Methods

# 2.1. Research Area

The research area of the Khairpur district is situated in the Southeast of Pakistan, bounded by Shikarpur and Sukkur districts in the North, Sanghar and Nawabshah districts in the South, with Larkana and the Indus River in the West and India to the East (Figure 1). The reason for selecting the Khairpur as a research area is that all the major orchards including bananas, dates and mangoes can be found in this region.



**Figure 1.** Geographic location of study area and spatial distribution of (**A**) validation dataset (**B**) training datasets.

The topography of Khairpur can be divided into a plain cultivated area in the Northern and Western parts of the district and desert area to the East and Southeast. This area is very rich and fertile in the vicinity of the Indus River and its irrigation canals. The Eastern part of Khairpur lies in the Nara desert, which is part of the great Indian desert. This area is characterized by extreme temperatures and severe droughts. The district has two well defined seasons: hot summers which begin after March and continue until October as well as cold winters with December, January and February as the coldest months. In addition, dust storms are frequent during April, May and June while the cooling Southern winds during summer nights are the only soothing element. Mean annual rainfall in this area is 120 to 230 mm mostly occurring during July and August. Based on its natural conditions, the orchards in this focused study area have been planted for decades; hence, making it one of the major fruit production areas in Pakistan. The primary orchards here include dates, mango and banana plantations, while cotton, wheat and sugarcane are the other main agricultural crops.

#### 2.2. Training and Validation Data

To train the ML algorithm for classification tasks, ground samples representing different LC types were needed. For this study, the training data were collected on a random samples basis, using android-based GPS application through a field survey conducted in late July 2022. Due to the harsh weather conditions and the high flood alert in this area during the survey period (July 2022), 70% of the samples were collected during the field survey while the remaining 30% were collected from high-resolution google earth images. The total 180 sample polygons representing four categories of LC (banana (50), dates (40), mango (30) and other LC (60)) were used as the training dataset (Figure 1B) for classification of LC. Other LC includes the mosaic of agriculture land, barren land, built-up area and water bodies. For accuracy assessment of the classified maps, overall, 625 systematic grid points were generated at a distance of  $200 \times 200$  m throughout the area and validated during the survey and from Google Earth high resolution imagery (Figure 1A).

## 2.3. Data Processing

#### 2.3.1. Sentinel-2 Data

In the present study, monthly based S2 Level-2 (L2A) images from January to December 2022, with cloud cover of less than 10% were downloaded from the website https://scihub.copernicus.eu/ assessed on 4 March 2023. L2A products have been pre-processed for atmospheric and geometric correction. S2 mission consists of the Sentinel 2A and 2B satellites equipped with a multispectral instrument, which provides global coverage at 5-days' time intervals [46].

The S2 has 13 bands with different spectral resolutions ranging from 0.433 to  $2.19\mu m$ . In visible and NIR regions, it has four 10 m bands. Six 20 m red-edge, NIR and SWIR bands as well and three 60 m bands for characterizing aerosol, cirrus cloud and water vapor [47]. In this study, 60 m bands including band-1, band-9 and band-10 were eliminated because these were specified for characterization of atmosphere and not for land surface observation. The 20 m spatial resolutions bands were resampled to 10 m spatial resolution utilizing an S2 resampling operator with nearest neighborhood techniques in SNAP 9.0 software.

#### 2.3.2. Vegetation Indices

To measure the vegetation growth and health conditions, remote sensing indices are the main indicators. In this study, seven vegetation indices were used, calculated for all multitemporal S2 images including Normalized Difference Vegetation Index (NDVI), Green Normalized Difference Vegetation Index (GNDVI), Transformed Normalized Difference Vegetation Index (TNDVI, Soil Adjusted Vegetation Index (SAVI), S2 Red-Edge Position (S2REP), Inverted Red-Edge Chlorophyll Index (IRECI) and Pigment Specific Simple Ratio (PSSRa). The formulas of these indices are given in Table 1.

Full Name	Index	Formula
Normalized Difference Vegetation Index	NDVI	<u>NIR – Red</u> NIR + Red
Green Normalized Difference Vegetation Index	GNDVI	$\frac{NIR - Green}{NIR + Green}$
Transformed Normalized Difference Vegetation Index	TNDVI	$\sqrt{\left(rac{NIR-Red}{NIR+Red}+0.5 ight)}$
Soil Adjusted Vegetation Index	SAVI	$(1+L)  imes rac{(NIR - Red)}{(NIR + Red + L)}$
S2 Red-Edge Position	S2REP	$705 + 35 \times ((\textit{Red} + \textit{NIR})/2 - \textit{Red2})/(\textit{Red3} - \textit{Red2}))$
Inverted Red-Edge Chlorophyll Index	IRECI	$(NIR - Red1) / \left(\frac{Red2}{Red3}\right)$
Pigment Specific Simple Ratio	PSSRa	<u>NIR</u> Red

Table 1. Vegetation indices derived from the Sentinel-2 images.

NDVI is the common vegetation index used to monitor vegetation density, intensity, health and growth [48]. It exploits the vigor and strength of the earth's surface vegetation. The GNDVI developed by [49] was more reliable than NDVI in identifying the different concentration rates of chlorophyll. A green spectral band was shown to be more effective than a red band for assessing vegetation chlorophyll variability [49]. TNDVI is the square root of NDVI and indicates the relationship between the amount of green biomass in a pixel [50]. The SAVI introduced by [51] aims to be a cross between the ratio-based indices and the perpendicular indices. A soil adjustment factor (L) is included in SAVI to distinguish the impact of soil brightness. The factor ranges from 0 to 1 where the maximum value is used in regions with more visible soil. For intermediate vegetation densities, 0.5 is the conventional value that is mostly used in most applications [52]. The S2REP index based on linear interpolation was presented by [53]. Red-edge, as the point of strong red absorption to near infrared reflectance, includes the information on growth status and chlorophyl content of crops. The reflectance at the inflexion point was determined and REP was obtained through interpolation of S2 red-edge band-5 and -6, respectively. The IRECI incorporates four spectral bands to estimate canopy chlorophyll content [54]. The PSSRa, introduced by [55], investigates the range potential of spectral bands for quantifying pigment at the scale of the whole plant canopy. It has the strongest and most linear relationships with canopy concentration per unit area of chlorophyll and carotenoids.

#### 2.3.3. Sentinel-1 Data

With the development of S1 mission, the SAR applications have gained increased usage due to the global data availability at free cost and their high temporal resolution. The S1 mission is a constellation of polar-orbiting Sentienl-1A and -1B satellites, prepared with C-band SAR sensors enabling data acquirement in all weather conditions [17]. S1 data are acquired in dual polarization with vertical transmit and horizontal receive (*VH*) and vertical transmit as well as (*VV*) backscattering, polarized vertical for transmission and horizontal/vertical for reception [56]. To examine the sensitivity of SAR backscatter for orchards classification, in the present study, multi-temporal S1 Level-1 Ground Range Detected (GRD) interferometric wide swath (IW) products were obtained with a temporal range of January to December 2022 (one image per month) from the website https://scihub.copernicus.eu/ assessed on 4 March 2023. The S1 images at GRD level are dual polarization *VH* and *VV* products which need to be preprocessed before being used for analysis. However, it required less processing time as compared to the Single Look Complex products.

The preprocessing of S1 GRD products includes four major steps: radiometric calibration, speckle filtering, radiometric terrain flattening and geometric correction. Radiometric calibration converts the pixels values to backscatter coefficients [57]. Before performing analysis in SAR, data speckle noise removal is necessary, as this degrades the quality of the image, making it difficult to understand the backscatter response from different LC [58]. To reduce

the speckle noise effect, we applied a Lee filter with  $3 \times 3$  window size to all multi-temporal images. The geometric correction of SAR images is also important because of the topographic variations in the scene and inclinations of the sensor, which distort the acquired images. As geometric correction is not considered in the GRD images, therefore, it is essential to perform and increase the Geolocation accuracy. The Range Doppler Terrain correction method was used for correction; we applied the Range Doppler Terrian Correction method by using a 30 m Copernicus Digital Elevation Model (DEM) to correct the topographic deformations. The Copernicus DEM was acquired automatically in SNAP software [59]. After the preprocessing, two polarization ratios were also calculated for all the multi-temporal images. The simple ratio of *VV* and *VH* (Equation (1)) and normalized ratio between the *VV* and *VH* bands (Equation (2)) [60] was calculated using the following formulas.

simple ratio 
$$=\frac{VH}{VV}$$
 (1)

normalized ratio 
$$=\frac{VV - VH}{VV + VH}$$
 (2)

# 2.4. Multicollinearity Test (MCT)

Utilization of MTMS data requires investigation of the correlation between MTMS variables. Using highly correlated variables can influence the model performance, using too many variables is also time consuming. Therefore, it was desirable to exclude the highly correlated variables and produce a model that can contribute to high accuracy and performance. In the present approach, based on the assumption that orchards may have the same spectral response at different time periods leads to correlation of variables, the MCT among all the multi-temporal S1 and S2 variables was performed using a psych package in R 2023.03 software, in order to eliminate the highly correlated variables. The analysis revealed that multi-temporal S2 spectral bands, vegetation indices and S1 bands and ratios were highly correlated. Hence, the threshold value of correlation >0.7, suggested in previous studies [34,35], was set to remove all the variables with a correlation greater than this threshold value. Figure 2 depicts variables used for classification after resolving the issue of multicollinearity.



Figure 2. Spearman correlation: Sentinel-1 and Sentinel-2 features with correlation < 0.7.

## 2.5. Feature Combination Schemes

The total multi-temporal S1 and S2 variables were 252. However, after the MCT, only 56 multi-temporal variables were left among which 13 were S2 spectral bands, 16 vegetation indices, 5 S1 bands and 22 S1 bands ratios. To quantify the multi-temporal S1 and S2, both individually and combined impacts on the classification of orchard types, six different feature combination schemes were developed. These schemes were used as predictors for the classification of banana, dates, mango and other LC, in the study area. The details of each scheme are given in Table 2.

Table 2. Feature combination schemes.

Schemes	Description	Number of Bands
Scheme-1	Spectral bands	13
Scheme-2	Spectral bands + Vegetation Indices	29
Scheme-3	SAR variables	27
Scheme-4	SAR variables + Vegetation Indices	43
Scheme-5	SAR variables + Spectral bands	40
Scheme-6	SAR variables + Spectral bands + Vegetation Indices	56

# 2.6. Spectral Separability Analysis

Spectral separability refers to quantifying the degree of separation between LC classes in different remote sensing images. Jeffries-Matusita (JM) distance is the indicator of separability commonly used in the applications of remote sensing [61–65]. JM distance was recommended to be more reliable for assessing separability [66] and demonstrated increased reliability for classes within homogenous distributions [67]. Therefore, we choose JM distance to examine the spectral separability among orchard types in multi-temporal S1 and S2 images.

JM distance [68] was calculated as

$$JM = 2\left(1 - e^{-B}\right) \tag{3}$$

where, Bhattacharyya distance is represented with B as given below:

$$B = \frac{1}{8}(x-y)^t \left(\frac{\Sigma_x + \Sigma_y}{2}\right)^{-1}(x-y) + \frac{1}{2}\ln\left(\frac{\left|\frac{\Sigma_x + \Sigma_y}{2}\right|}{\sqrt{|\Sigma_x| \cdot |\Sigma_y|}}\right)$$
(4)

where, x is the spectral response of the first object vector and y is the spectral response of the second object vector,  $\Sigma x$  and  $\Sigma y$  is the covariance matrix of samples x and y, alternatively.

JM distance value ranges from 0 to 2, where 0 value represents complete inseparability and a value of 2 represents completely separable [69]. We performed the JM distance analysis by using the training samples statistics at each class based on probability distribution.

#### 2.7. Classification Algorithms

A variety of classification algorithms are available for LC mapping, but the ML-based RF and SVM methods are widely recognized for their reliable performance across a range of remote sensing applications [70–73]. RF is an ensemble learning method based on decision trees [74,75]. It predicts the classification class by constructing numerous decision trees, each trained on randomly selected samples [76]. Final classification is determined by majority votes across all trees [77]. Two parameters, namely the number of predictors (mtry) and the number of trees (ntree), must be chosen before training the RF model. For the presented dataset, mtry was set to 56 and ntree to 500. SVM suggested by [78], is a ML method based on a statistical theory that performs nonlinear classification using kernel functions [79]. Input vectors are mapped into a high-dimensional feature space through

nonlinear transformation, and the SVM model identifies the best classification hyperplane to maximize the separation between classes in this space [80]. SVM primarily achieves fitting nonlinear functions by choosing a kernel function and other parameters, including gamma, which controls the smoothness of the hyperplane, and the cost, which controls the penalty for misclassification. In this study, we opted for the polynomial kernel function with gamma and cost values both set to 1.

## 2.8. Accuracy Assessment

To measure the performance of classification methods, accuracy assessment is very important. Commonly used accuracy assessment indicators are user accuracy (UA), producer accuracy (PA), overall accuracy (OA) and kappa coefficient (KC) [81,82], and are obtain through a confusion matrix [83]. A total of 625 systematic points, generated at a fixed distance of  $200 \times 200$  m throughout the study area (Figure 1B), were used to calculate the confusion matrix and obtain the accuracy assessment indicators.

#### 2.9. Feature Importance Measurement

Feature importance measurement aims to rank the relative importance of input multiple variables in ML models [84] and can be measured in several ways. The RF algorithms have benefits over other ML algorithms by having a built-in feature important measurement. The primary benefit of the RF method is that all the important input features can be calculated during the training of the RF model. In this study, Gini Index (GI) [85] was used for the feature importance measurement. The Gini index described the whole reduction in impurity of all nodes average and whole decision trees [86,87]. GI computes to reduce different samples through the whole nodes fragmented by a predictor variable in RF model. Gini examines the advantage of each feature, applying the information of all the features in the decision trees. However, all the features of the six classification schemes were input into the SVM without additional feature selection.

#### 3. Results

## 3.1. Spectral Separability Analysis

The JM distance between each pair of orchard classes were calculated for singledate, multi-temporal and fusion of S1 and S2 datasets. The detailed results are shown in (Figure 3a) representing three pairwise orchards combinations. The horizontal axes represent datasets and vertical time scale individually. The color of each grid signifies the JM distance value of the resultant datasets and time. The redder the grid cell, the higher the JM distance value and more independent the two classes at time scale specifically. Banana was highly distinguishable from dates in blue, red, red-edge bands, NIR bands, NDVI, TNDVI, GNDVI, SAVI, PSSRA and IRECI index during the summer months (May, June, July and August), because these are the peak growing months (vegetation density is high) and SWIR-1 and SWIR-2 from January and February. The single date S1 variables throughout the year were not found suitable for separation of banana from dates. Because of the degree of lushness during the summer months, the mango and banana were highly distinguished in Red-Edge and NIR bands, GNDVI, SAVI, IRECI and S2REP and also in Red, SWIR bands and all vegetation indices expect S2REP from the winter months (January, February). However, no separation was detected between mango and banana in blue, green bands and S1 variables throughout the year. Dates and mango were least distinguishable from each other in all the S1 and S2 variables except for the summer months blue band, as well as September, October and November NDVI and TNDVI. However, different SAR and optical variables have different capacities for vegetation discrimination. All in all, red-edge, NIR bands and vegetation indices were superior to the other spectral bands and S1 variables. During the peak growing period (May to Aug), red-edge and NIR bands were slightly better than the vegetation indices in distinguishing banana from the other orchards.



**Figure 3.** Spectral separability between orchards in (**a**) single-date images (**b**) multi-temporal images (**c**) fusion of optical and SAR variables.

Consequently, the JM distance of multi-temporal S1 and S2 variables was advance calculated as shown in (Figure 3b). The horizontal axes represent S1 and S2 variables, while the vertical axes represent the time scale in unit of two months, i.e., the two-month composite period. For the whole year, six temporal combinations were developed to calculate the JM distance. The results show that an increase in time series length increased the JM distance meaningfully. Once the time series length was greater than 4 months, banana-dates and mango-banana classes become almost completely separable, the JM distance reached 2.0, except for the simple ratio of VH/VV. However, the JM distance among dates-mango increased gradually with increasing the time series length. Once half year of time series composite was used, dates-mango were moderately separable. Furthermore, when the whole year composite was used, the JM distance reached 2.0 between dates-mango in blue band, red band, SWIR-1 band, NDVI, TNDVI and PSSRA only. For all the S1 variables, the dates-mango discrimination ability was lower even when the whole year composite was used. However, it can be observed that spectral bands and vegetation indices perform better than S1 variables. It is remarkable that for the most indistinguishable pair of dates-mango, multi-temporal S2 variables outperformed S1 variables.

The JM distance of the multi-temporal fusion of S1 and S2 features was also calculated, as shown in (Figure 3c). The results show that orchard classes were completely separable when multi-temporal optical and SAR features were combined while only date-mango was less separable in multi-temporal SAR features. In the fusion of multi-temporal S1 and S2 features, the JM distance value approached 2.0 for all the pairs of orchards including banana-dates, mango-banana and dates-mango. These results were in parallel with the classification results as, combining multi-temporal S1 and S2 features, achieved highest classification accuracy.

# 3.2. Assessment of Different Classification Schemes

Classification schemes with different features combinations were classified using RF and SVM methods. The overall accuracy and kappa value of each classification scheme are shown in (Figure 4). It was observed that RF and SVM algorithms achieved the highest accuracies of 93% and 94%, respectively, with scheme-6 which is the fusion of all multi-temporal S1 and S2 variables. Scheme-2 (optical bands + vegetation indices) ranked second with an accuracy of 91% and 93% for RF and SVM, respectively. Using only multi-temporal SAR variables (scheme-3) produces the relatively least reliable results with an accuracy of less than 80% in both RF and SVM methods. Compared with scheme-3, the combination of SAR variables with vegetation indices and optical bands (scheme-4 and scheme-5) significantly improved the accuracy of orchards classification. However, compared to other classification schemes, scheme-6 retrieved the best classification results in both RF and SVM methods. Hence, it revealed that a combination of multi-temporal optical and SAR features increased the classification accuracy.



Figure 4. Classification results with different feature combination schemes.

# 3.3. Comparison of the Classification Algorithms

The comparative analysis revealed that the SVM method slightly outperformed the RF method in terms of accuracy across all classification schemes as shown in Figure 4. Across scheme-1 to -6, the OA and KC for SVM were consistently higher compared to RF, except in scheme-4 where RF showed a slight advantage with a marginally higher OA and KC value by 1% and 0.7%, respectively.

The PA and UA of each orchard class, classified by using different classification schemes with RF and SVM algorithms, are shown in (Tables 3 and 4). The UA and PA of the mango class was relatively low with the SVM method and much lower with the RF method using multi-temporal SAR features only. In all other classification schemes, mango achieves greater than 70% UA and PA with RF and SVM methods, respectively. For the dates class, the highest PA with RF and SVM was reached by 96% and 99%, respectively, in scheme-6. However, the same approach produces the lower UA for dates as 82% and 87%, respectively, mainly due to confusion with mango class. The banana class with RF and SVM approaches achieves PA and UA higher than 90% in all the classification schemes except for scheme-3, while RF produced less than 90% PA and less than 86% UA in this scheme. Overall, the lowest UA (<70%) was obtained by mango and dates with both RF and SVM methods using scheme-3 (multi-temporal SAR features only). While in the same classification scheme,

banana achieved a good UA (>80%) with both classification methods. The UA and PA of mango, dates and banana were relatively high with both classification algorithms using scheme-6. These findings were consistent with the JM distance separability analysis (as discussed in Section 3.1). This is easy to comprehend because the JM distance for all the orchard classes approaches 2.0 when all the multi-temporal features (scheme-6) were used.

Random Forest (RF)						
Scheme-1	Others	Mango	Banana	Dates	User accuracy	
Others	353	1	2	1	98.88	
Mango	1	147	3	41	76.56	
Banana	7	2	187	1	94.92	
Dates	25	31	10	198	75	
Producer accuracy	91.45	81.22	92.57	82.16		
Scheme-2	Others	Mango	Banana	Dates	User accuracy	
Others	354	0	1	3	98.88	
Mango	6	154	5	3	91.67	
Banana	11	0	191	6	91.83	
Dates	15	27	5	229	82.97	
Producer accuracy	91.71	85.08	94.55	95.02		
Scheme-3	Others	Mango	Banana	Dates	User accuracy	
Others	320	14	4	51	82.26	
Mango	10	78	2	27	66.67	
Banana	6	9	186	19	84.54	
Dates	52	76	16	157	52.16	
Producer accuracy	82.47	44.07	89.42	61.81		
Scheme-4	Others	Mango	Banana	Dates	User accuracy	
Others	363	1	14	5	94.78	
Mango	2	143	0	6	94.70	
Banana	2	1	177	5	95.67	
Dates	19	36	11	225	77.32	
Producer accuracy	94.04	79.00	87.62	93.36		
Scheme-5	Others	Mango	Banana	Dates	User accuracy	
Others	372	0	0	3	99.2	
Mango	1	142	2	28	82.08	
Banana	0	0	186	1	99.46	
Dates	13	39	14	209	76	
Producer accuracy	96.37	78.45	92.08	86.72		
Scheme-6	Others	Mango	Banana	Dates	User accuracy	
Others	368	0	0	3	99.19	
Mango	4	150	1	5	93.75	
Banana	3	1	193	1	97.47	
Dates	11	30	8	232	82.56	
Producer accuracy	95.34	82.87	95.54	96.26		

 Table 3. Confusion matrix of RF with different feature combination schemes.

Support Vector Machine (SVM)						
Scheme-1	Others	Mango	Banana	Dates	User accuracy	
Others	357	0	3	0	99.17	
Mango	0	161	2	11	92.53	
Banana	2	1	188	0	98.43	
Dates	27	19	9	230	80.70	
Producer accuracy	92.49	88.96	93.07	95.44		
Scheme-2	Others	Mango	Banana	Dates	User accuracy	
Others	365	0	2	2	98.91	
Mango	1	163	4	14	89.56	
Banana	2	0	191	0	98.96	
Dates	18	18	5	225	84.59	
Producer accuracy	94.56	90.05	94.55	93.36		
Scheme-3	Others	Mango	Banana	Dates	User accuracy	
Others	324	2	1	28	91.27	
Mango	3	121	5	49	67.98	
Banana	1	1	190	4	96.94	
Dates	60	53	12	173	58.05	
Producer accuracy	83.50	68.36	91.35	68.11		
Scheme-4	Others	Mango	Banana	Dates	User accuracy	
Others	336	0	7	0	97.96	
Mango	0	159	0	12	92.98	
Banana	2	0	176	0	98.88	
Dates	48	22	19	229	72.01	
Producer accuracy	87.05	87.84	87.13	95.02		
Scheme-5	Others	Mango	Banana	Dates	User accuracy	
Others	357	1	3	0	98.89	
Mango	0	154	2	19	88	
Banana	2	1	190	0	98.44	
Dates	27	25	7	222	79.00	
Producer accuracy	92.49	85.08	94.06	92.12		
Scheme-6	Others	Mango	Banana	Dates	User accuracy	
Others	363	0	2	1	99.18	
Mango	4	166	5	1	94.32	
Banana	2	1	191	0	98.45	
Dates	17	14	4	239	87.23	
Producer accuracy	94.04	91.71	94.55	99.17		

Table 4. Confusion matrix of SVM with different feature combination schemes.

# 3.4. Overall Classification Results and Accuracy Assessment

The spatial distribution maps of the classification from different feature combination schemes using RF and SVM are depicted in (Figure 5). It can be spatially observed that dates are largely distributed in the East and Southeast of the research area while mangos are in the middle and Southeast of the area. Bananas cover a large area and are distributed

throughout the focused area. Other LCs, such as agricultural land, built-up areas and water bodies, collectively contribute a small amount of the total area. The classified maps obtained from the RF and SVM methods, using different features combinations, are consistent in most regions of the area. However, through the comparative analysis of the confusion matrixes results from RF and SVM methods (Tables 3 and 4), a high confusion of mango and dates was found in all the classification schemes which leads to the misclassification of two classes. The classification results of RF and SVM algorithms for banana and other classes are relatively consistent among all the feature combination schemes. In addition, the speckled effect observed over the LC classes with the RF method as shown in Figure 5(b1,c1,d1,f1) was resolved with the SVM method (Figure 5(b2,c2,d2,f2)). Moreover, the RF produced relatively little "salt and pepper" maps as compared to SVM maps in all the schemes (Figure 5). Overall findings illustrate the synthesis of multi-temporal SAR and show that optical features have a significant role in the LC classification improvement.



**Figure 5.** Classified maps with RF and SVM methods; a1, b1, c1, d1, e1, f1 and a2, b2, c2, d2, e2, f2 enlarged areas for comparative analysis.

# 3.5. Evaluating Features Importance for Orchards Classification

The feature significance scores were examined by the average reduction in the Gini Index (GI), using the RF model (Figure 6). Based on the top 20 important features in GI, multi-temporal SWIR and NIR bands, S2REP and NDVI indices are the most important features for orchards classification. Particularly, the multi-temporal S2REP index was more relevant, and in GI, the top 20 features has 7 multi-temporal S2REP indices, respectively.



Figure 6. Feature importance scores.

When assessing the relevance of time periods, March and August were the most important months for the discrimination of orchard classes as three features from each month were ranked in the top 20 important features. In addition, March to October also appeared to be the critical time period as there was at least one high ranked feature from each month. February was the least important month as no single feature was ranked in the top 20 important features. Among the multi-temporal S1 features, October-*VH* was among the top 15 essential features. However, all other features were ranked much lower in GI scores.

#### 4. Discussion

#### 4.1. Multi-Temporal Imagery

We applied two approaches to examine the orchards judgment capacity of multitemporal S1 and S2 datasets. In the first approach, JM distance was used to specify the separability among each orchard pair in single date images, multi-temporal and fusion of multi-temporal S1 and S2 images. The whole output illustrates that the multi-temporal data show superiority to single-date data. During the summer months (May to August), red-edge bands and NIR bands perform better to distinguish banana-dates and mangobanana. However, the performance for dates-mango in the same months was relatively poor. Date-mango only showed dissimilarity in the blue band during the growing season which may be linked to the blue band sensitivity to the water content in vegetation [88,89] and vegetation indices during monsoon season. In addition, the SWIR bands also showed good performance in January and February for discriminating banana-dates and mango-banana. These results correlate with the results obtained in other studies. Nelson and Immitzer et al. [90,91] reported that the NIR, SWIR and red-edge bands were among the most important for discriminating different LC. On the other hand, orchard types were difficult to differentiate from each other in all the single-date S1 variables. Also, the multi-temporal VH/VV ratio showed no advantage in discriminating orchard pairs on all dates. The VV and VH backscatter are more useful than VH/VV for discriminating orchard types [92]. Also, in the dates-mango discrimination, all the S1 multi-temporal variables showed almost no discrimination or very low JM distance value. Furthermore, the JM distance of the S1 and S2 variables mostly approached 2.0 when more than half year multi-temporal images were combined (Figure 3b). Hence, it can be concluded that S1 data could not be used alone due to the limited number of channels for discriminating the orchard types [93]. However, the fusion of multi-temporal S1 and S2 images demonstrate exceptional performance in the separation of orchards pairs. Similar to previous studies on fusion of S1 and S2 [94–96], from the combination of multi-temporal SAR and optical imageries, high classification accuracy was attained, compared with the result of multi-temporal SAR and optical alone.

In the second approach, we implemented RF and SVM algorithms to classify orchard categories in the study area by using different multi-temporal features combination. The best feature combination scheme was made up of a fusion of multi-temporal S1 and S2 features (scheme-6) with an overall accuracy of 93% and 94% with SVM and RF methods, correspondingly. This indicates that multi-temporal remote sensing imagery complements the information in single-date images and enhances the predictive power of the classification models. Moreover, these outcomes support previous studies concerning LC classification using multi-temporal data [90,97,98]. Hill et al. [99] utilized multi-temporal high-resolution Airborne Thematic Mapper images for the classification of the deciduous tree species. It was found that a single image provides satisfactory results with an accuracy of 71%, but combined multi-dates images notably achieves a higher accuracy of 84%. Kollert et al. [100] applied multi-temporal S2 images for the classification of tree species and their output illustrated that classification output attained with S2 temporal images was 10% higher than the single S2 images. Additionally, Zhu and Liu [101] achieved the highest accuracy of 88% for forest type classification using dense time series imagery. In our research, the multi-temporal optical bands (scheme-1) performed well independently but the overall accuracy plateaued when additional multi-temporal features were combined. Notably, this could be due to the spectral variation in orchard types being different in different time periods. Moreover, the multi-temporal SAR images (scheme-3) poorly performed, likely due to the small variations in backscatter coefficients of the orchard types. However, together with either optical bands or vegetation indices, it made up good classification results. These findings were similar to the findings of [102], who assessed the potential of S1 alone, S2 alone and combination of both S1 and S2 data for mapping forest-agricultural mosaic over Cantabrian Range, Spain and Paragominas in Brazil. The results obtained from S2 alone were more accurate than S1 alone. However, the highest accuracy was achieved by the fusion of S1 and S2. This may be due to the fact that SAR sensors are sensitive to vegetation structure while optical sensors are sensitive to the concentration of vegetation chlorophyll. Consequently, upon examination of a 10m spatial resolution, orchard classes can be easily distinguished based on their physiology and physical structure. Patel et al. [92] found that both polarizations VV and VH and their ratio VH/VV contributed equally but to a low importance in classifying the forest-agricultural mosaic. Thus, this leads to the conclusion that only the C-band SAR dataset is not appropriate to classify orchard types. Generally, C-band SAR has a faster single saturation rate and lower penetration depth than

L-band SAR, it is less effective for monitoring trees [103]. Furthermore, Patel et al. [92] also highlighted that L-band is more sensitive to a variation in plant density than the C-band, which interacts mostly with the primary branches of the canopy, while the L-band is within the vegetation canopy. However, in the case of bare soil and artificial surfaces, the SAR data sometimes discriminate better than the optical data as the SAR single reacts differently to these two LC classes [104].

#### 4.2. Performance on Orchard Types

In this study, three major orchard types in the focused study area including dates, mango and banana were classified with higher accuracy through the fusion of S1 and S2 multi-temporal data. Among these orchards, mango-dates were highly confused with each other in all the multi-temporal feature combination schemes. This was probably due to the reason that these two orchards mostly have a heterogenous pattern inside the study area shown in (Figure 5). Mango has the lowest PA of 44% and 68% in SAR data only, using RF and SVM, respectively. Similarly, using SAR data only, the dates retrieved the lowest accuracy of 61% and 68% with RF and SVM, respectively. While for both mango and dates, the PA reached 91% and 99% in a combination of multi-temporal SAR and optical datasets. As noticed, RF showed a substantial underestimation of orchard types, which produced low PA as relate to SVM method. The difference in the confusion matrices, from scheme-1 to scheme-6, is that the PA of mango and dates greatly improved. The reason is that the combination of multi-temporal S1 and S2 data increased the two classes separability. As compared to mango and dates, banana was less confused with these and vice versa in all the combination schemes. It could be due to the fact that bananas have a homogenous pattern in the research area. The research findings indicated that in similar prospective studies, bananas can be separated by using multi-temporal S1 and S2 data only. The class specific UA and PA is reported in (Tables 3 and 4). However, for illustrative purposes, thematic maps (Figure 5) have also been included. For similar prospective studies, thematic mapping focusing on this study area may suggestively incorporate more LC classes, such as agricultural crop types, barren land and built-up area. The outputs of the present research are encouraging; homogenous areas of orchard types instead of mixed tree species, which vary greatly on spectral basis and are generally more challenging to classify [105] have been explored.

#### 4.3. Feature Importance

In the present study, the MCT excluded 196 of the 252 features, presumably because they were highly correlated ( $R^2 > 0.7$ ) with the 56 features that were selected. Among these 56 features, the temporal SWIR band, NDVI and S2REP indices were ranked highly in the Gini Index, as well as typically from the growing month which strengthened the position of these images. Furthermore, a majority of their S2REP indices ranked among the top 10 features. However, discrepancies occur regarding the importance of specific features. The study's results from [90,91] reports that the blue and red bands were more critical. However, in this study, these bands were removed due to their high correlation with other bands.

#### 4.4. Performance of Classification Algorithms

In this study, the classification performance of RF and SVM algorithms was evaluated. These algorithms demonstrate less sensitivity to the choice of input features [18] and are equally reliable [106–108]. The maps in (Figure 5), obtained through RF and SVM methods with different features combinations, were quite similar, except for differences in accuracies. Indeed, the kappa values and overall accuracy revealed discrepancies in the output classifications. Overall, SVM performance is superior to the RF algorithm which is consistent with previous studies [109,110]. However, other research studies comparing SVM and RF have reported contrasting results [111–113] possibly due to the application of a small training dataset. The RF model used a huge number of decision trees for the

training data of high dimensional tree selection [85]; hence, it is crucial to train the RF model with a small dataset [114]. The SVM model can escape a small training dataset, as well as high-dimensional decision trees selection [79]. Moreover, other studies of SVM and RF comparisons propose the advantages of SVM method [115–117]. The details of the kappa coefficient and the overall accuracy attained by the SVM and RF methods using different classification schemes is shown in (Tables 3 and 4).

#### 5. Conclusions

This research examined the effectiveness of S1 and S2 multi-temporal SAR and optical data for orchard classification using two ML algorithms in the sub-arid region of Pakistan. The use of multi-temporal SAR data proved unreliable for the orchard discrimination, due to insufficient spectral dissimilarity. Though the multi-temporal S2 optical data performed relatively better than the SAR data, the fusion of multi-temporal S1 and S2 showed that high classification accuracy can be achieved regarding the accuracies with each sensor independently. In addition, the spectral dissimilarity in the fused datasets is high as compared to the single-date or temporal S1 and S2 datasets. Particularly, for the most indistinguishable pair dates-mango, fused data achieve a relatively high JM distance value and overall accuracy. However, the results of the present research suggested most multi-temporal variables are highly correlated; hence, there is need to conduct the MCT and remove highly correlated variables to reduce overfitting, as well as the time needed for analysis.

Although SVM was superior to RF in orchard classification, the differences in their performance were not significant, suggesting RF can also be considered a viable option. The important features ranked by the Gini Index were multi-temporal SWIR bands, NIR bands, S2REP and NDVI indices. Considering the S1 bands and ratios derived from these bands, multi-temporal S2 perform best in the orchard discrimination. The findings demonstrated that the multi-temporal fusion of SAR and optical is well-suited for orchard mapping using ML methods and can be applied to future research such as vegetation mapping, change detection and monitoring in the heterogeneous landscapes. The outcomes provide baseline information for the spatial-distribution-based mapping of orchards, yield estimations and agricultural practices. Future research should consider advanced deep learning and fusion techniques for improving LC classification and include more LC classes for detailed spatial distribution-based maps.

**Author Contributions:** Conceptualization, A.U.R. and L.Z.; methodology, A.U.R. and L.Z.; resources, L.Z.; data collection and software, A.U.R. and M.M.S., supervision, L.Z.; validation, M.M.S. and A.R.; writing, review and editing, A.U.R., M.M.S., L.Z. and A.R.; revision of original manuscript, A.U.R. and A.R. All authors have read and agreed to the published version of this manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant number 41830108).

Data Availability Statement: Readers can contact authors for availability of data and materials.

Acknowledgments: The authors acknowledge the Chinese Academy of Sciences and the Alliance of International Science Organizations (ANSO) for awarding the ANSO Scholarship for Young Talents to Arif UR Rehman to carry out this research project.

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

# References

- Panda, S.S.; Hoogenboom, G.; Paz, J.O. Remote sensing and geospatial technological applications for site-specific management of fruit and nut crops: A review. *Remote Sens.* 2010, 2, 1973–1997. [CrossRef]
- Chen, B.; Xiao, X.; Wu, Z.; Yun, T.; Kou, W.; Ye, H.; Lin, Q.; Doughty, R.; Dong, J.; Ma, J. Identifying establishment year and pre-conversion land cover of rubber plantations on Hainan Island, China using landsat data during 1987–2015. *Remote Sens.* 2018, 10, 1240. [CrossRef]
- Altman, J.; Doležal, J.; Čížek, L. Age estimation of large trees: New method based on partial increment core tested on an example of veteran oaks. For. Ecol. Manag. 2016, 380, 82–89. [CrossRef]

- 4. Tan, C.-W.; Wang, D.-L.; Zhou, J.; Du, Y.; Luo, M.; Zhang, Y.-J.; Guo, W.-S. Assessment of Fv/Fm absorbed by wheat canopies employing in-situ hyperspectral vegetation indexes. *Sci. Rep.* **2018**, *8*, 9525. [CrossRef] [PubMed]
- Bolton, D.K.; Friedl, M.A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. For. Meteorol.* 2013, 173, 74–84. [CrossRef]
- 6. McMorrow, J. Relation of oil palm spectral response to stand age. Int. J. Remote Sens. 1995, 16, 3203–3209. [CrossRef]
- Buddenbaum, H.; Schlerf, M.; Hill, J. Classification of coniferous tree species and age classes using hyperspectral data and geostatistical methods. *Int. J. Remote Sens.* 2005, 26, 5453–5465. [CrossRef]
- Vuolo, F.; Neuwirth, M.; Immitzer, M.; Atzberger, C.; Ng, W.-T. How much does multi-temporal Sentinel-2 data improve crop type classification? *Int. J. Appl. Earth Obs. Geoinf.* 2018, 72, 122–130. [CrossRef]
- 9. Zhang, J. Multi-source remote sensing data fusion: Status and trends. Int. J. Image Data Fusion 2010, 1, 5–24. [CrossRef]
- Gómez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogramm. Remote Sens.* 2016, 116, 55–72. [CrossRef]
- Esch, T.; Metz, A.; Marconcini, M.; Keil, M. Combined use of multi-seasonal high and medium resolution satellite imagery for parcel-related mapping of cropland and grassland. *Int. J. Appl. Earth Obs. Geoinf.* 2014, 28, 230–237. [CrossRef]
- 12. Peña, M.; Brenning, A. Assessing fruit-tree crop classification from Landsat-8 time series for the Maipo Valley, Chile. *Remote Sens. Environ.* **2015**, *171*, 234–244. [CrossRef]
- Dostálová, A.; Wagner, W.; Milenković, M.; Hollaus, M. Annual seasonality in Sentinel-1 signal for forest mapping and forest type classification. *Int. J. Remote Sens.* 2018, 39, 7738–7760. [CrossRef]
- 14. Persson, M.; Lindberg, E.; Reese, H. Tree species classification with multi-temporal Sentinel-2 data. *Remote Sens.* **2018**, *10*, 1794. [CrossRef]
- 15. Immitzer, M.; Neuwirth, M.; Böck, S.; Brenner, H.; Vuolo, F.; Atzberger, C. Optimal input features for tree species classification in Central Europe based on multi-temporal Sentinel-2 data. *Remote Sens.* **2019**, *11*, 2599. [CrossRef]
- Zhou, X.-X.; Li, Y.-Y.; Luo, Y.-K.; Sun, Y.-W.; Su, Y.-J.; Tan, C.-W.; Liu, Y.-J. Research on remote sensing classification of fruit trees based on Sentinel-2 multi-temporal imageries. *Sci. Rep.* 2022, *12*, 11549. [CrossRef]
- 17. Dobrinić, D.; Gašparović, M.; Medak, D. Sentinel-1 and 2 time-series for vegetation mapping using random forest classification: A case study of Northern Croatia. *Remote Sens.* **2021**, *13*, 2321. [CrossRef]
- Schulz, D.; Yin, H.; Tischbein, B.; Verleysdonk, S.; Adamou, R.; Kumar, N. Land use mapping using Sentinel-1 and Sentinel-2 time series in a heterogeneous landscape in Niger, Sahel. *ISPRS J. Photogramm. Remote Sens.* 2021, 178, 97–111. [CrossRef]
- 19. Inglada, J.; Vincent, A.; Arias, M.; Marais-Sicre, C. Improved early crop type identification by joint use of high temporal resolution SAR and optical image time series. *Remote Sens.* **2016**, *8*, 362. [CrossRef]
- Pham, L.H.; Pham, L.T.; Dang, T.D.; Tran, D.D.; Dinh, T.Q. Application of Sentinel-1 data in mapping land-use and land cover in a complex seasonal landscape: A case study in coastal area of Vietnamese Mekong Delta. *Geocarto Int.* 2022, 37, 3743–3760. [CrossRef]
- Kpienbaareh, D.; Sun, X.; Wang, J.; Luginaah, I.; Bezner Kerr, R.; Lupafya, E.; Dakishoni, L. Crop type and land cover mapping in northern Malawi using the integration of sentinel-1, sentinel-2, and planetscope satellite data. *Remote Sens.* 2021, 13, 700. [CrossRef]
- 22. Steinhausen, M.J.; Wagner, P.D.; Narasimhan, B.; Waske, B. Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 595–604. [CrossRef]
- 23. Gerstl, S. Physics concepts of optical and radar reflectance signatures A summary review. *Int. J. Remote Sens.* **1990**, *11*, 1109–1117. [CrossRef]
- Nhu, V.-H.; Mohammadi, A.; Shahabi, H.; Ahmad, B.B.; Al-Ansari, N.; Shirzadi, A.; Geertsema, M.; R. Kress, V.; Karimzadeh, S.; Valizadeh Kamran, K. Landslide detection and susceptibility modeling on cameron highlands (Malaysia): A comparison between random forest, logistic regression and logistic model tree algorithms. *Forests* 2020, *11*, 830. [CrossRef]
- 25. Forkuor, G.; Dimobe, K.; Serme, I.; Tondoh, J.E. Landsat-8 vs. Sentinel-2: Examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. *GIScience Remote Sens.* **2018**, *55*, 331–354. [CrossRef]
- Dalponte, M.; Bruzzone, L.; Gianelle, D. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sens. Environ.* 2012, 123, 258–270. [CrossRef]
- 27. Dalponte, M.; Ørka, H.O.; Gobakken, T.; Gianelle, D.; Næsset, E. Tree species classification in boreal forests with hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 2632–2645. [CrossRef]
- George, R.; Padalia, H.; Kushwaha, S.P.S. Forest tree species discrimination in western Himalaya using EO-1 Hyperion. Int. J. Appl. Earth Obs. Geoinf. 2014, 28, 140–149. [CrossRef]
- Prospere, K.; McLaren, K.; Wilson, B. Plant species discrimination in a tropical wetland using in situ hyperspectral data. *Remote Sens.* 2014, 6, 8494–8523. [CrossRef]
- Baldeck, C.A.; Asner, G.P.; Martin, R.E.; Anderson, C.B.; Knapp, D.E.; Kellner, J.R.; Wright, S.J. Operational tree species mapping in a diverse tropical forest with airborne imaging spectroscopy. *PloS One* 2015, 10, e0118403. [CrossRef]
- Georganos, S.; Grippa, T.; Vanhuysse, S.; Lennert, M.; Shimoni, M.; Kalogirou, S.; Wolff, E. Less is more: Optimizing classification performance through feature selection in a very-high-resolution remote sensing object-based urban application. *GIScience Remote Sens.* 2018, 55, 221–242. [CrossRef]

- 32. Leroux, L.; Jolivot, A.; Bégué, A.; Seen, D.L.; Zoungrana, B. How reliable is the MODIS land cover product for crop mapping sub-Saharan agricultural landscapes? *Remote Sens.* **2014**, *6*, 8541–8564. [CrossRef]
- Peña, J.M.; Gutiérrez, P.A.; Hervás-Martínez, C.; Six, J.; Plant, R.E.; López-Granados, F. Object-based image classification of summer crops with machine learning methods. *Remote Sens.* 2014, 6, 5019–5041. [CrossRef]
- Ur Rehman, A.; Ullah, S.; Shafique, M.; Khan, M.S.; Badshah, M.T.; Liu, Q.-j. Combining Landsat-8 spectral bands with ancillary variables for land cover classification in mountainous terrains of northern Pakistan. J. Mt. Sci. 2021, 18, 2388–2401. [CrossRef]
- 35. Bobrowski, M.; Gerlitz, L.; Schickhoff, U. Modelling the potential distribution of Betula utilis in the Himalaya. *Glob. Ecol. Conserv.* **2017**, *11*, 69–83. [CrossRef]
- 36. Irfana Noor, M.; Sanaullah, N.; Barkat Ali, L. Economic efficiency of banana production under contract farming in Sindh Pakistan. J. Glob. Econ. 2015, 3, 2.
- 37. Dahri, G.N.; Talpur, B.A.; Nangraj, G.M.; Mangan, T.; Channa, M.H.; Jarwar, I.A.; Sial, M. Impact of climate change on banana based cropping pattern in District Thatta, Sindh Province of Pakistan. *J. Econ. Impact* **2020**, *2*, 103–109. [CrossRef]
- 38. Usman, M.; Fatima, B.; Khan, M.M.; Chaudhry, M.I. Mango in Pakistan: Achronological Review. *Pak. J. Agric. Sci.* 2003, 40, 151–154.
- 39. Badar, H.; Ariyawardana, A.; Collins, R. Dynamics of mango value chains in Pakistan. Pak. J. Agric. Sci. 2019, 56, 523–530.
- 40. Hussain, D.; Butt, T.; Hassan, M.; Asif, J. Analyzing the role of agricultural extension services in mango production and marketing with special reference to world trade organization (WTO) in district Multan. *J. Agric. Soc. Sci.* **2010**, *6*, 6–10.
- 41. Mehdi, M.; Ahmad, B.; Yaseen, A.; Adeel, A.; Sayyed, N. A comparative study of traditional versus best practices mango value chain. *Pak. J. Agric. Sci.* **2016**, *53*.
- 42. Abul-Soad, A.A.; Mahdi, S.M.; Markhand, G.S. Date Palm Status and Perspective in Pakistan. Date Palm Genetic Resources and Utilization: Volume 2: Asia and Europe; Springer: Berlin/Heidelberg, Germany, 2015; pp. 153–205.
- 43. Kousar, R.; Sadaf, T.; Makhdum, M.S.A.; Iqbal, M.A.; Ullah, R. Competiveness of Pakistan's selected fruits in the world market. *Sarhad J. Agric.* 2019, 35, 1175–1184. [CrossRef]
- Memon, M.I.N.; Noonari, S.; Kalwar, A.M.; Sial, S.A. Performance of date palm production under contract farming in Khairpur Sindh Pakistan. J. Biol. Agric. Healthc. 2015, 5, 19–27.
- 45. Khushk, A.M.; Memon, A.; Aujla, K.M. Marketing channels and margins of dates in Sindh, Pakistan. J. Agric. Res. 2009, 47, 293–308.
- 46. Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.R.; Murayama, Y.; Ranagalage, M. Sentinel-2 data for land cover/use mapping: A review. *Remote Sens.* **2020**, *12*, 2291. [CrossRef]
- 47. Wang, Q.; Shi, W.; Li, Z.; Atkinson, P.M. Fusion of Sentinel-2 images. Remote Sens. Environ. 2016, 187, 241–252. [CrossRef]
- Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 1979, *8*, 127–150. [CrossRef]
- 49. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]
- 50. Senseman, G.M.; Bagley, C.F.; Tweddale, S.A. Correlation of rangeland cover measures to satellite-imagery-derived vegetation indices. *Geocarto Int.* **1996**, *11*, 29–38. [CrossRef]
- 51. Huete, A.R. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 1988, 25, 295–309. [CrossRef]
- 52. Muller, S.J.; Sithole, P.; Singels, A.; Van Niekerk, A. Assessing the fidelity of Landsat-based fAPAR models in two diverse sugarcane growing regions. *Comput. Electron. Agric.* **2020**, *170*, 105248. [CrossRef]
- 53. Guyot, G.; Baret, F. Utilisation de la haute resolution spectrale pour suivre l'etat des couverts vegetaux. In Proceedings of the Spectral Signatures of Objects in Remote Sensing, Aussois, France, 18–22 January 1988; p. 279.
- Clevers, J.; De Jong, S.; Epema, G.; Addink, E.; Van Der Meer, F.; Skidmore, A. Meris and the Red-edge index. In Proceedings of the Second EARSeL Workshop on Imaging Spectroscopy, Enschede, The Netherlands, 11–13 July 2000.
- 55. Blackburn, G.A. Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *Int. J. Remote Sens.* **1998**, *19*, 657–675. [CrossRef]
- Li, L.; Kong, Q.; Wang, P.; Xun, L.; Wang, L.; Xu, L.; Zhao, Z. Precise identification of maize in the North China Plain based on Sentinel-1A SAR time series data. *Int. J. Remote Sens.* 2019, 40, 1996–2013. [CrossRef]
- 57. Kaplan, G.; Avdan, U. Monthly analysis of wetlands dynamics using remote sensing data. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 411. [CrossRef]
- Adiri, Z.; El Harti, A.; Jellouli, A.; Lhissou, R.; Maacha, L.; Azmi, M.; Zouhair, M.; Bachaoui, E.M. Comparison of Landsat-8, ASTER and Sentinel 1 satellite remote sensing data in automatic lineaments extraction: A case study of Sidi Flah-Bouskour inlier, Moroccan Anti Atlas. *Adv. Space Res.* 2017, 60, 2355–2367. [CrossRef]
- Twele, A.; Cao, W.; Plank, S.; Martinis, S. Sentinel-1-based flood mapping: A fully automated processing chain. *Int. J. Remote Sens.* 2016, 37, 2990–3004. [CrossRef]
- 60. Filgueiras, R.; Mantovani, E.C.; Althoff, D.; Fernandes Filho, E.I.; Cunha, F.F.d. Crop NDVI monitoring based on sentinel 1. *Remote Sens.* **2019**, *11*, 1441. [CrossRef]
- 61. Liu, H.; Zhang, F.; Zhang, L.; Lin, Y.; Wang, S.; Xie, Y. UNVI-based time series for vegetation discrimination using separability analysis and random forest classification. *Remote Sens.* **2020**, *12*, 529. [CrossRef]

- 62. Wardlow, B.D.; Egbert, S.L.; Kastens, J.H. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains. *Remote Sens. Environ.* **2007**, *108*, 290–310. [CrossRef]
- 63. Arvor, D.; Jonathan, M.; Meirelles, M.S.P.; Dubreuil, V.; Durieux, L. Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. *Int. J. Remote Sens.* 2011, *32*, 7847–7871. [CrossRef]
- 64. Yeom, J.; Han, Y.; Kim, Y. Separability analysis and classification of rice fields using KOMPSAT-2 High Resolution Satellite Imagery. *Res. J. Chem. Environ.* **2013**, *17*, 136–144.
- 65. Swain, P.; Robertson, T.; Wacker, A. Comparison of the divergence and B-distance in feature selection. *LARS Inf. Note* **1971**, 20871, 41399–47906.
- 66. Swain, P.H.; Davis, S.M. Remote sensing: The quantitative approach. *IEEE Trans. Pattern Anal. Mach. Intell.* **1981**, *3*, 713–714. [CrossRef]
- 67. Thomas, I.; Ching, N.; Benning, V.; D'aguanno, J. Review Article A review of multi-channel indices of class separability. *Int. J. Remote Sens.* **1987**, *8*, 331–350. [CrossRef]
- 68. Richards, J.A.; Richards, J.A. Remote sensing digital image analysis; Springer: Berlin/Heidelberg, Germany, 2022; Volume 5.
- 69. Jensen, J.R. Digital Image Processing: A Remote Sensing Perspective; Sprentice Hall: Upper Saddle River, NJ, USA, 2005.
- Yin, Q.; Liu, M.; Cheng, J.; Ke, Y.; Chen, X. Mapping paddy rice planting area in northeastern China using spatiotemporal data fusion and phenology-based method. *Remote Sens.* 2019, *11*, 1699. [CrossRef]
- Zhang, C.; Mishra, D.R.; Pennings, S.C. Mapping salt marsh soil properties using imaging spectroscopy. *ISPRS J. Photogramm. Remote Sens.* 2019, 148, 221–234. [CrossRef]
- 72. Chen, L.; Wang, Y.; Ren, C.; Zhang, B.; Wang, Z. Optimal combination of predictors and algorithms for forest above-ground biomass mapping from Sentinel and SRTM data. *Remote Sens.* **2019**, *11*, 414. [CrossRef]
- 73. Blomley, R.; Hovi, A.; Weinmann, M.; Hinz, S.; Korpela, I.; Jutzi, B. Tree species classification using within crown localization of waveform LiDAR attributes. *ISPRS J. Photogramm. Remote Sens.* **2017**, *133*, 142–156. [CrossRef]
- 74. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. Classification and Regression Trees; Taylor & Francis Group: Abingdon, UK, 1984.
- 75. Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* 2002, *2*, 18–22.
- 76. Pal, M. Random forest classifier for remote sensing classification. Int. J. Remote Sens. 2005, 26, 217–222. [CrossRef]
- 77. Xi, Y.; Ren, C.; Tian, Q.; Ren, Y.; Dong, X.; Zhang, Z. Exploitation of time series sentinel-2 data and different machine learning algorithms for detailed tree species classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7589–7603. [CrossRef]
- 78. Vapnik, V.N. An overview of statistical learning theory. IEEE Trans. Neural Netw. 1999, 10, 988–999. [CrossRef] [PubMed]
- 79. Cortes, C.; Vapnik, V. Support vector machine. Mach. Learn. 1995, 20, 273–297. [CrossRef]
- 80. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. Stat. Comput. 2004, 14, 199–222. [CrossRef]
- 81. Friedl, M.A.; Brodley, C.E. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **1997**, *61*, 399–409. [CrossRef]
- Rwanga, S.S.; Ndambuki, J.M. Accuracy assessment of land use/land cover classification using remote sensing and GIS. Int. J. Geosci. 2017, 8, 611. [CrossRef]
- 83. Stehman, S.V. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* **1997**, *62*, 77–89. [CrossRef]
- 84. Friedman, J.H. Greedy function approximation: A gradient boosting machine. Ann. Stat. 2001, 1189–1232. [CrossRef]
- 85. Breiman, L. Random Forests. Mach. Learn. 2001, 5–32. [CrossRef]
- 86. Nembrini, S.; König, I.R.; Wright, M.N. The revival of the Gini importance? Bioinformatics 2018, 34, 3711–3718. [CrossRef]
- 87. Boulesteix, A.-L.; Bender, A.; Lorenzo Bermejo, J.; Strobl, C. Random forest Gini importance favours SNPs with large minor allele frequency: Impact, sources and recommendations. *Brief. Bioinform.* **2012**, *13*, 292–304. [CrossRef]
- 88. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
- Jackson, T.J.; Chen, D.; Cosh, M.; Li, F.; Anderson, M.; Walthall, C.; Doriaswamy, P.; Hunt, E.R. Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. *Remote Sens. Environ.* 2004, 92, 475–482. [CrossRef]
- Nelson, M. Evaluating Multitemporal Sentinel-2 Data for Forest Mapping Using Random Forest. Master's Thesis, Stockholm University, Stockholm, Sweden, 2017.
- 91. Immitzer, M.; Vuolo, F.; Atzberger, C. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens.* **2016**, *8*, 166. [CrossRef]
- Patel, P.; Srivastava, H.S.; Panigrahy, S.; Parihar, J.S. Comparative evaluation of the sensitivity of multi-polarized multi-frequency SAR backscatter to plant density. *Int. J. Remote Sens.* 2006, 27, 293–305. [CrossRef]
- Zeyada, H.H.; Ezz, M.M.; Nasr, A.H.; Shokr, M.; Harb, H.M. Evaluation of the discrimination capability of full polarimetric SAR data for crop classification. *Int. J. Remote Sens.* 2016, 37, 2585–2603. [CrossRef]
- Sun, L.; Chen, J.; Han, Y. Joint use of time series Sentinel-1 and Sentinel-2 imagery for cotton field mapping in heterogeneous cultivated areas of Xinjiang, China. In Proceedings of the 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Istanbul, Turkey, 16–19 July 2019; pp. 1–4.
- Clerici, N.; Valbuena Calderón, C.A.; Posada, J.M. Fusion of Sentinel-1A and Sentinel-2A data for land cover mapping: A case study in the lower Magdalena region, Colombia. J. Maps 2017, 13, 718–726. [CrossRef]

- Mercier, A.; Betbeder, J.; Baudry, J.; Le Roux, V.; Spicher, F.; Lacoux, J.; Roger, D.; Hubert-Moy, L. Evaluation of Sentinel-1 & 2 time series for predicting wheat and rapeseed phenological stages. *ISPRS J. Photogramm. Remote Sens.* 2020, 163, 231–256.
- Reese, H.M.; Lillesand, T.M.; Nagel, D.E.; Stewart, J.S.; Goldmann, R.A.; Simmons, T.E.; Chipman, J.W.; Tessar, P.A. Statewide land cover derived from multiseasonal Landsat TM data: A retrospective of the WISCLAND project. *Remote Sens. Environ.* 2002, 82, 224–237. [CrossRef]
- 98. Wolter, P.T.; Mladenoff, D.J.; Host, G.E.; Crow, T.R. Using multi-temporal landsat imagery. *Photogramm. Eng. Remote Sens* **1995**, *61*, 1129–1143.
- 99. Hill, R.; Wilson, A.; George, M.; Hinsley, S. Mapping tree species in temperate deciduous woodland using time-series multispectral data. *Appl. Veg. Sci.* 2010, *13*, 86–99. [CrossRef]
- Kollert, A.; Bremer, M.; Löw, M.; Rutzinger, M. Exploring the potential of land surface phenology and seasonal cloud free composites of one year of Sentinel-2 imagery for tree species mapping in a mountainous region. *Int. J. Appl. Earth Obs. Geoinf.* 2021, 94, 102208. [CrossRef]
- Zhu, X.; Liu, D. Accurate mapping of forest types using dense seasonal landsat time-series. *ISPRS J. Photogramm. Remote Sens.* 2014, 96, 1–11. [CrossRef]
- 102. Mercier, A.; Betbeder, J.; Rumiano, F.; Baudry, J.; Gond, V.; Blanc, L.; Bourgoin, C.; Cornu, G.; Ciudad, C.; Marchamalo, M. Evaluation of Sentinel-1 and 2 time series for land cover classification of forest–agriculture mosaics in temperate and tropical landscapes. *Remote Sens.* 2019, *11*, 979. [CrossRef]
- 103. Woodhouse, I.H. Introduction to Microwave Remote Sensing; CRC press: Boca Raton, FL, USA, 2017.
- 104. Lee, J.-S.; Pottier, E. Polarimetric Radar Imaging: From Basics to Applications; CRC Press: Boca Raton, FL, USA, 2017.
- 105. Reese, H.; Nilsson, M.; Pahlén, T.G.; Hagner, O.; Joyce, S.; Tingelöf, U.; Egberth, M.; Olsson, H. Countrywide estimates of forest variables using satellite data and field data from the national forest inventory. *AMBIO: A J. Hum. Environ.* 2003, 32, 542–548. [CrossRef]
- Belgiu, M.; Drăgu, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* 2016, 114, 24–31. [CrossRef]
- 107. Ghosh, A.; Fassnacht, F.E.; Joshi, P.K.; Koch, B. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 49–63. [CrossRef]
- 108. Sesnie, S.E.; Finegan, B.; Gessler, P.E.; Thessler, S.; Ramos Bendana, Z.; Smith, A.M. The multispectral separability of Costa Rican rainforest types with support vector machines and Random Forest decision trees. *Int. J. Remote Sens.* 2010, *31*, 2885–2909. [CrossRef]
- Ghosh, A.; Joshi, P.K. A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. *Int. J. Appl. Earth Obs. Geoinf.* 2014, 26, 298–311. [CrossRef]
- Yuan, H.; Yang, G.; Li, C.; Wang, Y.; Liu, J.; Yu, H.; Feng, H.; Xu, B.; Zhao, X.; Yang, X. Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: Analysis of RF, ANN, and SVM regression models. *Remote Sens.* 2017, 9, 309. [CrossRef]
- Han, Z.; Zhu, X.; Fang, X.; Wang, Z.; Wang, L.; Zhao, G.-X.; Jiang, Y. Hyperspectral estimation of apple tree canopy LAI based on SVM and RF regression. Spectrosc. Spectr. Anal. 2016, 36, 800–805.
- 112. You, H.; Huang, Y.; Qin, Z.; Chen, J.; Liu, Y. Forest Tree Species Classification Based on Sentinel-2 Images and Auxiliary Data. *Forests* **2022**, *13*, 1416. [CrossRef]
- 113. Dahhani, S.; Raji, M.; Hakdaoui, M.; Lhissou, R. Land cover mapping using sentinel-1 time-series data and machine-learning classifiers in agricultural sub-saharan landscape. *Remote Sens.* **2022**, *15*, 65. [CrossRef]
- Deng, H.; Runger, G.; Tuv, E. Bias of importance measures for multi-valued attributes and solutions. In Proceedings of the Artificial Neural Networks and Machine Learning–ICANN 2011: 21st International Conference on Artificial Neural Networks, Proceedings, Part II 21. Espoo, Finland, 14–17 June 2011; pp. 293–300.
- 115. Ballanti, L.; Blesius, L.; Hines, E.; Kruse, B. Tree species classification using hyperspectral imagery: A comparison of two classifiers. *Remote Sens.* **2016**, *8*, 445. [CrossRef]
- 116. Sheykhmousa, M.; Mahdianpari, M.; Ghanbari, H.; Mohammadimanesh, F.; Ghamisi, P.; Homayouni, S. Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2020, 13, 6308–6325. [CrossRef]
- 117. Khatami, R.; Mountrakis, G.; Stehman, S.V. A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research. *Remote Sens. Environ.* 2016, 177, 89–100. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.