



Article

Machine Learning Techniques for Phenology Assessment of Sugarcane Using Conjunctive SAR and Optical Data

Md Yeasin ^{1,*} , Dipanwita Haldar ², Suresh Kumar ², Ranjit Kumar Paul ¹ and Sonaka Ghosh ³ ¹ ICAR-Indian Agricultural Statistics Research Institute, New Delhi 110012, India; ranjit.paul@icar.gov.in² Indian Institute of Remote Sensing, Uttarakhand 248001, India; dipanwita@iirs.gov.in (D.H.); suresh_kumar@iirs.gov.in (S.K.)³ ICAR-Research Complex for Eastern Region, Patna 800014, India; sonaka.ghosh@icar.gov.in

* Correspondence: md.yeasin@icar.gov.in

Abstract: Crop phenology monitoring is a necessary action for precision agriculture. Sentinel-1 and Sentinel-2 satellites provide us with the opportunity to monitor crop phenology at a high spatial resolution with high accuracy. The main objective of this study was to examine the potential of the Sentinel-1 and Sentinel-2 data and their combination for monitoring sugarcane phenological stages and evaluate the temporal behaviour of Sentinel-1 parameters and Sentinel-2 indices. Seven machine learning models, namely logistic regression, decision tree, random forest, artificial neural network, support vector machine, naïve Bayes, and fuzzy rule based systems, were implemented, and their predictive performance was compared. Accuracy, precision, specificity, sensitivity or recall, F score, area under curve of receiver operating characteristic and kappa value were used as performance metrics. The research was carried out in the Indo-Gangetic alluvial plains in the districts of Hisar and Jind, Haryana, India. The Sentinel-1 backscatters and parameters VV, alpha and anisotropy and, among Sentinel-2 indices, normalized difference vegetation index and weighted difference vegetation index were found to be the most important features for predicting sugarcane phenology. The accuracy of models ranged from 40 to 60%, 56 to 84% and 76 to 88% for Sentinel-1 data, Sentinel-2 data and combined data, respectively. Area under the ROC curve and kappa values also supported the supremacy of the combined use of Sentinel-1 and Sentinel-2 data. This study infers that combined Sentinel-1 and Sentinel-2 data are more efficient in predicting sugarcane phenology than Sentinel-1 and Sentinel-2 alone.

Keywords: machine learning; optical data; phenology; SAR data; sugarcane

Citation: Yeasin, M.; Haldar, D.; Kumar, S.; Paul, R.K.; Ghosh, S. Machine Learning Techniques for Phenology Assessment of Sugarcane Using Conjunctive SAR and Optical Data. *Remote Sens.* **2022**, *14*, 3249. <https://doi.org/10.3390/rs14143249>

Academic Editors: Peter Hofmann and Hossein M. Rizeei

Received: 10 May 2022

Accepted: 14 June 2022

Published: 6 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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/>).

1. Introduction

Globally, sugarcane is the most cultivated crop, producing approximately 1890 million tonnes on approximately 26.8 million hectares [1]. About 80% of the world's sugar is produced from sugarcane in tropical and subtropical climates [2]. India is the largest consumer and the second largest sugar producer in the world after Brazil [3]. Identification and assessment of phenological stages of crop growth are important for precision agriculture. Phenology is the analysis of the timing of biological events, the causes of their timing concerning biotic and abiotic forces, and the interrelationship between phases [4]. Crop phenology estimation of individual fields could provide vital inputs to support monitoring of farm productivity in smallholder systems, providing the required information to prove targeted approaches to improve agricultural resilience to climate change as well as livelihood security of the farmers [5,6]. Four phenology phases characterize the phenological stages of sugarcane, namely, germination (up to 40 days after planting), tillering (up to 120 days after planting), grand growth (up to 270 days after planting), and ripening phase or maturity. In the maturity phase, synthesis and accumulation of sugar take place [7,8]. The consistent and systematic remote monitoring of crop phenological dynamics is critical

for optimizing management activities of the farm and assessing agricultural resilience to extreme weather events targeting towards future climate change [9].

Remote sensing offers the unique opportunity to generate biophysical parameters required for crop management [10]. Because it provides precise and reliable information on vegetation development, remote sensing is very effective in monitoring phenology stages. Several studies on crop monitoring based on remotely sensed data have been conducted over the last two decades. Shihua et al. [11] tried to monitor rice phenology using Savitzky–Golay filter and wavelet transformation on enhanced vegetation index (EVI) derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data. Wei et al. [12] developed normalized difference vegetation index (NDVI) time-series reconstruction techniques using MODIS data for assessing crop phenology. Liu et al. [13] developed a new algorithm to assess real-time maize and soybean phenology using Visible Infrared Imaging Radiometer Suite (VIIRS) and MODIS data. Ghaderpour and Vujadinovic [14] suggested a method for detecting disturbances in vegetation time series using spectral and wavelet analysis in MATLAB and Python [15,16]. Sakamoto et al. [17,18] monitored soybean and maize phenological stages by applying the shape model to the wide dynamic range vegetation index (WDRVI) from MODIS data.

Sentinel-1 and Sentinel-2 data have gained significant popularity in recent years due to their high spatial and temporal resolution and wide availability. The optical data from Sentinel-2 and the synthetic aperture radar (SAR) polarimetric data from Sentinel-1 are extremely sensitive to crop phenological stages [19]. Song and Wang et al. [20] found that Sentinel-1 backscatter polarisation (VH and VV; V for vertical, H for horizontal) has a strong relationship with the vegetation canopy of rice and suggested its use for phenological study. Mercier et al. [21] studied the potential of the Sentinel-1 and Sentinel-2 data for monitoring wheat and rapeseed crop phenology. They inferred the supremacy of combined Sentinel-1 and Sentinel-2 data over their singleton use [22,23]. Song et al. [24] introduced the fusion of MODIS and Landsat/Sentinel-2 data for modelling crop phenology with higher accuracy. Haldar et al. [25] used dual-pol Shannon entropy and Radar Vegetation Index (RVI) from polarimetric Sentinel-1 for capturing mustard and wheat phenology. They reported a significant correlation between the crop stages and Shannon entropy. Chen et al. [26] (rice), Mercier et al. [21] (wheat and rapeseed), Narin and Abdikan [27] (sunflower), and Haldar et al. [28] (cotton) are some examples of the acceptance of Sentinel data for monitoring crop phenology.

In this study, we examined the sensitivity of various optical and SAR parameters to the phenological stages of sugarcane and found the important parameters to characterize the sugarcane phenology. The main objective of this study was to model sugarcane crop phenology using Sentinel-1 and Sentinel-2 data and compare the efficiency of these two varied data sources. For this purpose, seven machine learning algorithms, namely logistic regression, decision tree (DT), random forest (RF), artificial neural network (ANN), support vector machine (SVM), naïve Bayes and fuzzy rule based systems (FRBS) classification models were implemented, and their accuracy compared.

The remaining manuscript is structured as follows: Data description and pre-processing of satellite images and methodology are provided in Section 2. The results of the proposed model and temporal analysis of SAR and optical data are illustrated in Section 3. Discussions of the proposed method and its performance are presented in Section 4. Finally, the conclusions of this study are provided in Section 5. A list of abbreviations has been given in “Abbreviations”.

2. Materials and Methods

2.1. Study Region

The study was carried out in the Indo-Gangetic alluvial plains located in the Hisar and Jind districts of Haryana, India (Figure 1). The climate in the region is continental, with hot summers and mild winters [29]. Sugarcane, one of the major crops of Hisar, comprises at least 0.97 thousand ha of area under sugarcane cultivation [30].

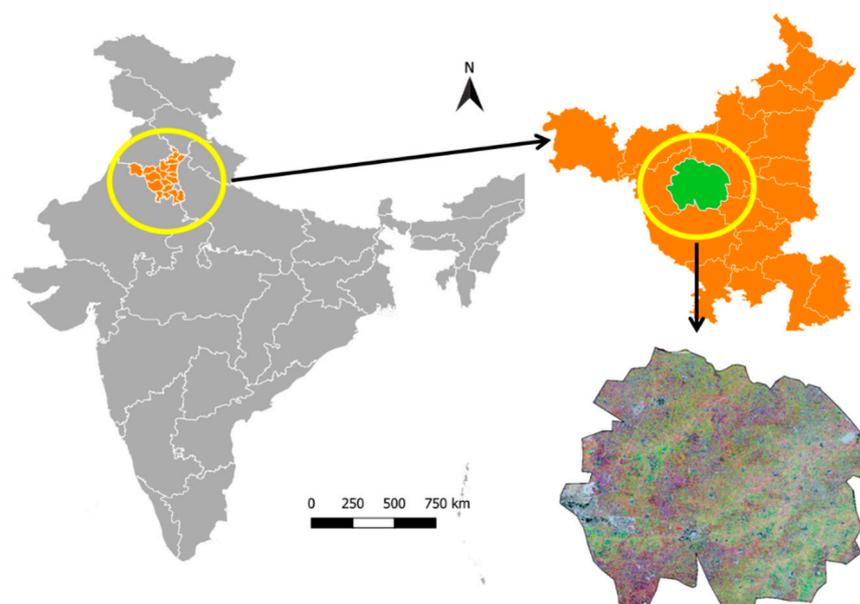


Figure 1. Study area.

Field data were collected from March 2020 to December 2020. To select ground truth points, a purposive sampling methodology was utilized. Purposive sampling was implemented by taking into account the size of the field (>2 ha), years of agriculture (>2 years), sowing time, variety, and accessibility. A total of 40 fields were selected, and phenological development was recorded temporally by multi-temporal observations conducted in each field. In our study area, the sugarcane crops are planted at three different times: 1st week of March, 1st week of April, and 1st week of May. So, they attain phenological stages at different times of the calendar year. It should be mentioned that during May and July 2020, we were unable to collect Sentinel-2 data from a few sampling points due to high cloud cover over the region. Figure 2 depicts different dates of acquisition with their recorded phenological stages.

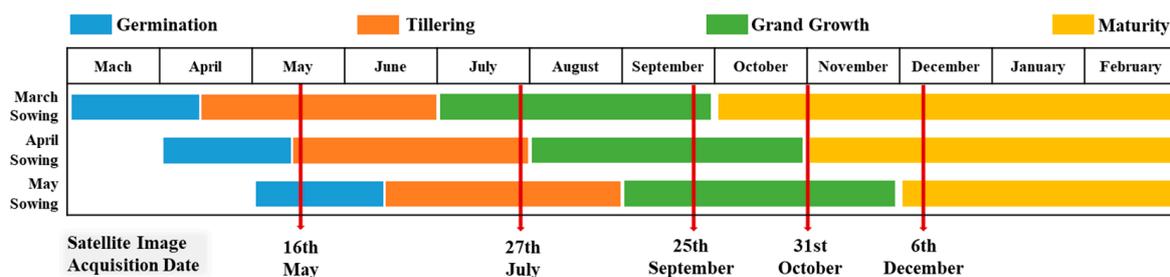


Figure 2. Satellite image acquisition.

2.2. Datasets

For the spatial-temporal analysis, the study used Sentinel-1 and Sentinel-2 datasets. A series of five Sentinel-1 and Sentinel-2 datasets were collected from May 2020 to December 2020 (16th May, 27th July, 25th September, 31st October, and 6th December) from the European Space Agency data hub (Copernicus Open Access Hub). Single Look Complex (SLC) Sentinel-1 data of VH-VV polarization with range resolution of 2.3 m and spatial resolution of 13.9 m were acquired. Sentinel-2 MSI (Multi-Spectral Instrument) data with level 2A image, (atmospherically and geometrically corrected) with spatial and spectral resolutions of 10–60 m and 13 m, respectively, were collected. A total of ten Sentinel-2, five Sentinel-1 SLC and five Sentinel-1 Ground Range Detected (GRD) images were collected and studied.

2.2.1. Pre-Processing of SLC Data

The Sentinel Application Platform (SNAP 8.0) was used to undertake all of the Sentinel-1 data processing steps (<https://step.esa.int/main/snap-8-0-released/> accessed on 5 May 2022). First, Sentinel-1 images were split into different sub swaths, and precise orbit files were applied and calibrated to a complex output. After calibration, Sentinel-1 images were co-registered, deburst, and merged. After that, a C_2 polarimetric matrix was generated, and a multi-looking and polarimetric speckle filter (Refined Lee with window size 5×5) was applied to reduce salt and pepper noise and increase image clarity [31]. Following that, images were decomposed into polarimetric parameters: entropy, anisotropy, and alpha angle using the H- α dual polarimetric decomposition algorithm with a window size of 5×5 . Finally, Range-Doppler terrain correction was applied with the help of the Shuttle Radar Topography Mission (SRTM) 3 arc second digital elevation model [32]. The final spatial resolution of the resulting parameters was $10 \text{ m} \times 10 \text{ m}$.

2.2.2. Pre-Processing of GRD Data

To obtain backscatter coefficients from Sentinel-1 images, SLC images were converted into GRD images performing a series of steps (thermal noise removal, calibration, TOPSAR deburst, multilook, speckle filter) in the SNAP 8.0 platform. A backscattering ratio was derived by dividing σ^0_{VH} by σ^0_{VV} . Using the following equation, the backscattering coefficients σ^0_{VH} and σ^0_{VV} , as well as the backscattering ratio $\sigma^0_{VH}:\sigma^0_{VV}$, were translated into decibels (db) [33].

$$\sigma^0(dB) = 10 \times \log_{10}(\sigma^0) \quad (1)$$

2.2.3. Pre-Processing of Sentinel-2 Data

Sentinel-2 MSI level 2A images were processed versions (atmospherically and geometrically corrected) of Sentinel-2 MSI level 1C images [34]. All bands of Sentinel-2 MSI level 2A images were resampled into the highest resolution (10 m) using the nearest-neighbour algorithm, and required indices were generated.

Based on extensive study of the literature, four indices, namely, the normalized difference vegetation index (NDVI), normalized difference water index (NDWI), weighted difference vegetation index (WDVI), and Sentinel-2 Red-Edge Position Index (S2REP), were computed from Sentinel-2 data using SNAP v8.0 software to study sugarcane phenology. NDVI was chosen because of its sensitivity to green vegetation [35,36] and crop vigour [37]. The NDWI was considered because of its responsiveness to water molecules in crop stems and leaves and also its short-wave infrared (SWIR) feature [38–40]. The WDVI was selected because it is associated with the canopy's chlorophyll content and is also used to determine LAI [41]. The S2REP was considered since it is a refined form of REP for Sentinel-2 utilising linear interpolation [42,43]. Details of Sentinel-2 indices are depicted in Table 1.

Table 1. Details of Sentinel-2 indices.

Index	Formula	Sentinel-2	Range	References
NDVI	$(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$	$(\text{Band } 8 - \text{Band } 4) / (\text{Band } 8 + \text{Band } 4)$	−1 to 1	[44]
NDWI	$(860 \text{ nm} - 1240 \text{ nm}) / (860 \text{ nm} + 1240 \text{ nm})$	$(\text{Band } 3 - \text{Band } 8) / (\text{Band } 3 + \text{Band } 8)$	−1 to 1	[38,45]
WDVI	$\text{NIR} - 0.5 * \text{RED}$	$\text{Band } 9 - 0.5 * \text{Band } 5$	−1 to 1	[46]
S2REP	$\frac{705 + 358 * ((\text{RED} + \text{VNIR}3) / 2 - \text{VNIR})}{(\text{VNIR}2 - \text{VNIR})}$	$\frac{705 + 358 * (((\text{Band } 4 + \text{Band } 7) / 2 - \text{Band } 5) * (\text{Band } 6 - \text{Band } 5))}{(\text{Band } 6 - \text{Band } 5)}$	650 to 750	[42]

2.3. Methods

Seven machine learning classifiers were used to perform classification on Sentinel-1 parameters, Sentinel-2 indices, and the composite of Sentinel-1 and Sentinel-2 data to identify sugarcane phenological phases. Ground data were collected on 40 plots at 5 temporal points. The number of virtual plots totalled 200. The datasets were normalised

using a min-max normalisation technique and randomly split into training and validation data in 80:20 ratios. Training and testing plots of one temporal point are presented in Figure 3.

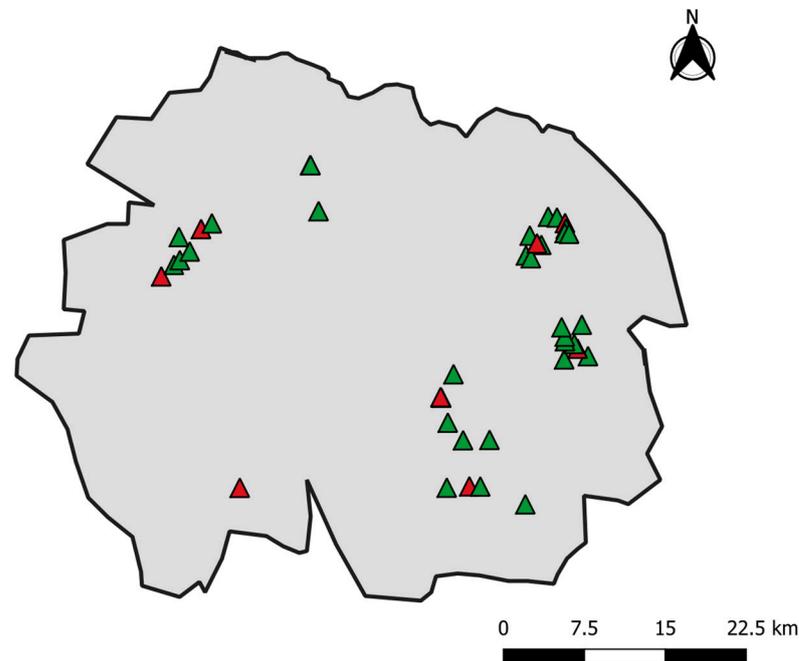


Figure 3. Training and testing plots of single temporal point.

All of the machine learning models, as mentioned earlier, were implemented on the training data, and performance metrics were calculated. In the literature, many performance metrics were found, but seven most commonly used metrics were used to evaluate the performance of the above-mentioned classification algorithms: accuracy, precision, specificity, sensitivity or recall, F score, area under curve of receiver operating characteristic (ROC) and kappa value [47,48]. Brief descriptions of the techniques used are presented in subsequent sections.

2.3.1. Logistic Regression

When the responses of a study are categorical in nature, the logistic regression model, introduced primarily by Cox [49] and Walker and Duncan [50], is more suited than the traditional regression model. The main objective of the logistic regression model is to model categorical variables or the probability of membership function of response variables regressed on one or more regression variables. Logistic regression, a supervised classification algorithm, is used to analyse the association between metric and non-metric variables. Multinomial logistic regression is used when there are more than two categories of response variable present. Multinomial logistic regression is widely applied due to its loose assumptions for homoscedasticity, normality and linearity [51]. Multinomial logistic regression can be represented by the following equation:

$$Y_i = \beta + \beta_1 x_1 + \dots + \beta_p x_p + \varepsilon_i \quad (2)$$

In our case, Y_i has four categories: germination, tillering, grand growth and maturity, which are represented by 0, 1, 2, and 3, respectively. The number of independent variables, p , were six, four and ten for Sentinel-1, Sentinel-2, and their combined data, respectively.

Let π_i denote the probability that $Y_i = 1$ given $X = (1, x_1, x_2, \dots, x_p)$. The logistic function is used to draw the relation between the probability π_i and X . The logistic function is an S-shaped curve represented by Equation (3).

$$\pi_i = \frac{1}{1 + e^{-(\beta + \beta_1 x_1 + \dots + \beta_p x_p)}}; -\infty < (\beta + \beta_1 x_1 + \dots + \beta_p x_p) < \infty \quad (3)$$

2.3.2. Naïve Bayes

Naïve Bayes is a probability classifier based on the Bayes theorem. Naive Bayesian networks (NB) are composed of directed acyclic graphs with only one parent (representing the unobserved node) and several children (corresponding to observed nodes). It has a rigid independence assumption between input variables [52]. The naïve Bayes classifier calculates conditional probability from likelihood and prior distribution using the Bayes theorem:

$$p(C_k | x_1, x_2, \dots, x_p) = \frac{p(C_k) p(x_1, x_2, \dots, x_p | C_k)}{p(x)} \quad (4)$$

where C_k is the classes, k is the number of classes. $p(C_k | x_1, x_2, \dots, x_p)$, $p(C_k)$, $p(x_1, x_2, \dots, x_p | C_k)$ and $p(x)$ are the posterior probability, prior probability, likelihood function, and total probability, respectively.

2.3.3. Support Vector Machine Learning

SVM learning is a learning theory-based machine learning method that was proposed by Vapnik [53]. SVM is one of the most efficient and user-friendly algorithms for solving classification problems [54]. Data points that are closest to the decision plane are called support vectors [55]. SVM aims to form a decision boundary that minimizes the classification error by maximizing margin size. For non-linear complex problems, the feature space is mapped to higher dimension space using a kernel function such that the new space is linearly separable. This technique is called kernel trick [56]. This enables the SVM to deal with high dimensional and complex problems efficiently.

2.3.4. Decision Tree

A decision tree is a learning algorithm with a tree-like structure [57]. It is made up of three elements: decision nodes, leaf nodes, and a root node. It separates a training dataset into branches, which are then subdivided into subbranches. This procedure is repeated until a leaf node is obtained [58]. The leaf node cannot be further divided. Decision nodes help to connect the leaves. The root node of the tree is the feature that best classifies the training data. It may be determined by information gain [59] and Gini index [60] criteria.

2.3.5. Random Forest

A random forest is a supervised learning technique derived from decision tree algorithms. It uses ensemble learning, which is a technique that combines multiple classifiers to solve complicated problems. The random forest algorithm's 'forest' is trained with bagging aggregation [61]. Bagging is an ensemble algorithm that increases the accuracy of machine learning models. It generates output by aggregating the output of several trees. The precision of the model improves as the number of trees grows. While growing the trees, it adds more randomness to the algorithms. Unlike decision tree, it does not search the most important feature for splitting a node; it finds the best feature among a random subset of features [62]. This improves the model accuracy as well as increasing the applicability of the model.

2.3.6. Neural Network

Because of their flexible design, neural networks are a popular machine learning technology that may be used to simulate a wide range of applications. In general, neural network architecture is composed of three layers (input layer, hidden layer, and output

layer); however, there are a wide range of designs available in the literature [63]. The number of neurons in the input layer or the number of neurons in the hidden layer may also be adjusted. Neural networks are based on three key elements: the input unit and activation functions, network architecture, and the weight of each unit. The logistic function is most popular as an activation function, given by

$$g(x) = \frac{1}{1 + \exp(-x)} \quad (5)$$

The first two aspects are decided before model fitting, and the weights need to be trained. There are several algorithms with which this network can be trained [64]. Back-propagation, a supervised learning algorithm, is the most widely used learning algorithm for neural networks [65].

2.3.7. FRBS (Fuzzy Rule Based Systems)

Zadeh [66] introduced fuzzy set theory as an extension of classical set theory to model based on degrees of membership. Using a membership function, the fuzzy system maps the crisp value to a fuzzy value, called fuzzification. Defuzzification is the process of restoring a fuzzy value to its original form. FRBS is a rule-based classification algorithm based on the fuzzy set theory. This algorithm uses “IF A THEN B” condition where A and B are fuzzy sets. A and B are the antecedent and consequent component, respectively. In general, the FRBS model consists of four steps: fuzzification, structure identification, parameter estimation and defuzzification [67,68]. After fuzzification, the rule base corresponding to pairs of input and output variables is determined in the second steps. Then, the membership function parameters are optimised and estimated, and finally, output values are de-fuzzified to produce the final output. A variety of FRBS models have been presented in the literature. Ishibuchi and Nakashima’s [69] model was adopted in this study.

3. Results

3.1. Temporal Analysis of Sentinel-1 Parameters and Sentinel-2 Indices

The above discussed (Section 2.2) four indices from Sentinel-2 and six parameters from Sentinel-1 were estimated and analysed. In this section, we sought to visualise the temporal pattern of these parameters and indices (Figures 4 and 5).

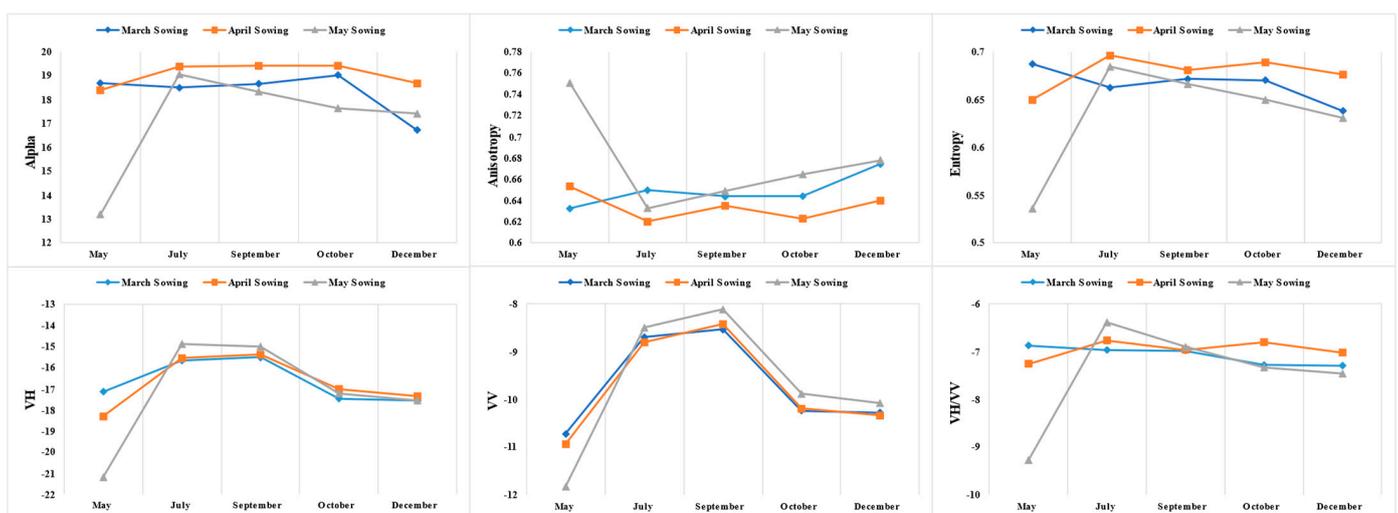


Figure 4. Temporal profile of Sentinel-1 parameters.

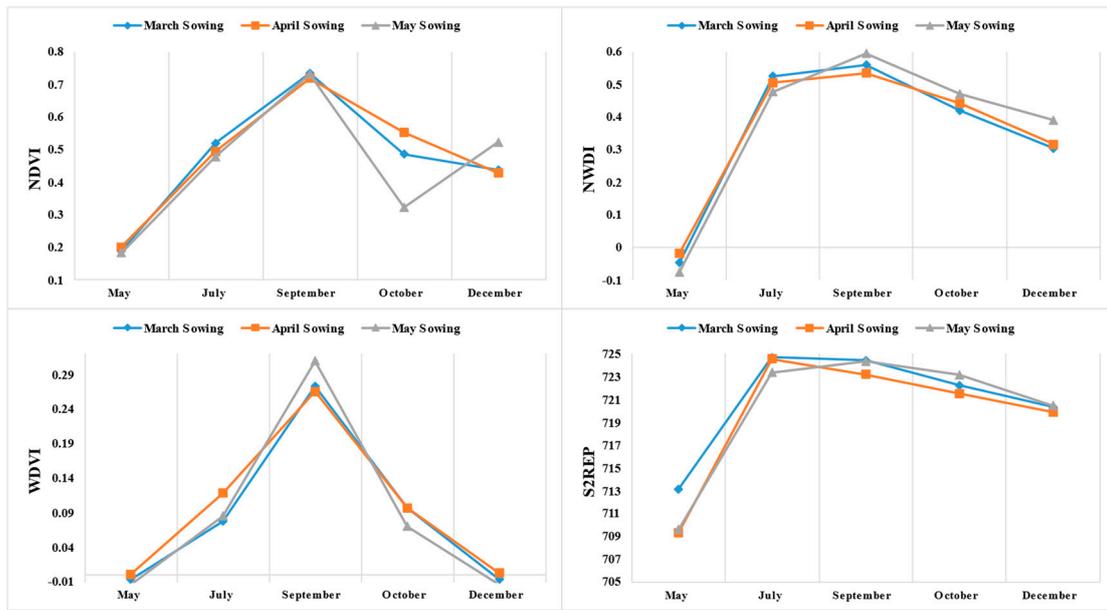


Figure 5. Temporal profile of Sentinel-2 indices.

The above Figures 4 and 5 show the behaviour of Sentinel-1 parameters and Sentinel-2 indices in different months for different sowing periods. By comparing the behaviour of Sentinel-1 parameters and Sentinel-2 indices with different image acquisition dates and other months, it is clear that Sentinel-2 indices were more responsive than Sentinel-2 parameters to the months and Sentinel-1 parameters were more responsive than Sentinel-2 indices to different sowing dates.

In Figures 6 and 7, we present the mean responses of Sentinel-1 and Sentinel-2 features in different phenological stages to visualise their behavior in different phenological stages.

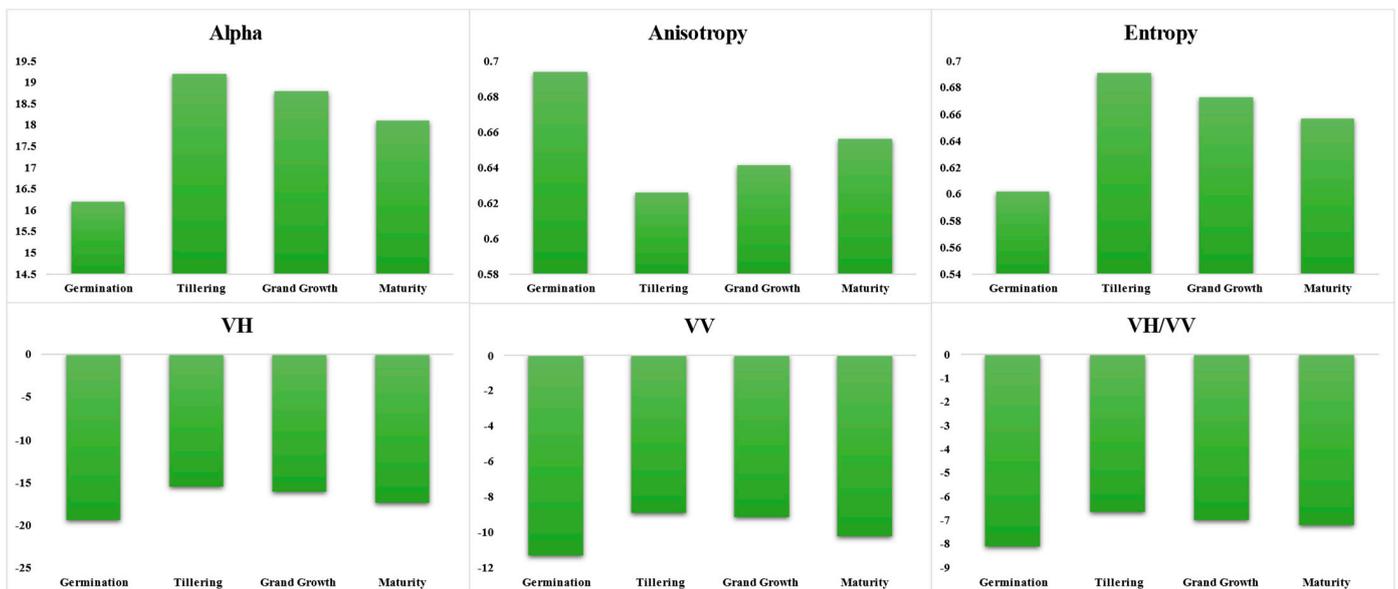


Figure 6. Response of Sentinel-1 parameters with respect to phenology stages.

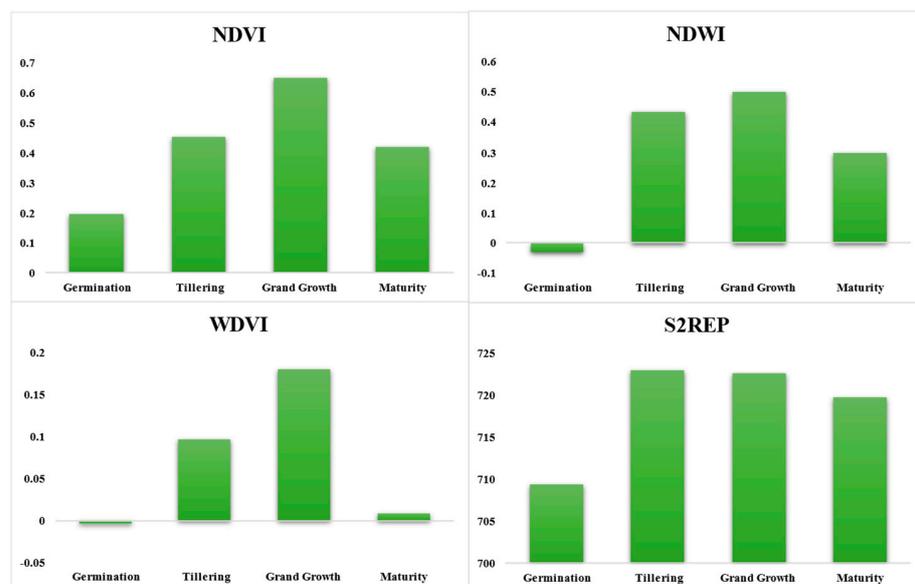


Figure 7. Response of Sentinel-2 indices with respect to phenology stages.

The behaviours of Sentinel-1 parameters and Sentinel-2 indices in various phenology stages are listed in Figures 6 and 7. The inferences drawn from the above Figure 6 are discussed below.

The alpha angle and entropy have similar behaviour in respective phenology stages and temporal profiles (Figures 4 and 6). They are lowest in May, increase in July, remain constant until October, and then decrease in December (Figure 4). At the tillering stage, they are at their highest, and at the germination stage, they are at their lowest (Figure 6). Anisotropy is maximum during the germination period, then quickly declines in the tillering stage, then gradually increases until maturity (Figure 4). It is highest in May, rapidly decreases in July, and increases until the end (Figure 6). Alpha increased in value to tillering owing to a rise in dominance scattering, i.e., volume, double bounce, or surface scattering as plant height increased and with the formation of new leaves [70,71]. Entropy also increased in tillering stages due to an increase in the randomness of the crops [72]. Alpha and entropy started decreasing from the grand growth to maturity stage due to partial lower leaf fall with thickening of the stems and juice accumulation and decline in vegetative vigour. Anisotropy decreased in the tillering stage due to high first dominance scattering (i.e., volume scattering) as the increase in homogeneousness in vegetation [73]. A slight rise in anisotropy was observed in grand growth to maturity due to evidence of two dominance scattering as stem enlargement occurred in sugarcane.

The VV started with the lowest absolute value in May, rapidly increased to the highest value in July and September, then steadily decreased in October and December. The VH showed behavior similar to that of VV, but VV saturated earlier than VH and was more dynamically sensitive to the later stage of sugarcane. VH/VV started with the lowest value in May, suddenly increased in July, then constantly decreased, as seen in Figure 4. All GRD parameters (VV, VH, and VH/VV) were identical in response to various phenological stages (Figure 6). They attained the lowest value during germination and exhibited slight variation. The absolute value of the VH backscatter was always greater than that of VV backscatter. VV backscatter was more responsive to vegetation growth than VH backscatter, as it surged from July to September, whilst VH showed a constant increase [74]. VV is sensitive to vegetation moisture, as demonstrated by Fieuzal et al. [75] and Cookmartin et al. [76], during the tillering and grand growth phases. These are dominated by the surface soil moisture, roughness, soil texture, and row orientation/geometry [77]. Because of the impact of soil and crop height, low values of backscatter coefficients in germination were recorded. These SAR backscatter coefficients rose during tillering due to an increase in the double bounce between the vertical stalks and the soil [13,78] and during the grand growth

phase due to an increase in volume scattering due to rapid accumulation of aboveground biomass [78,79]. As the vegetation vigour on the sugarcane canopy declined in maturity, the backscatter coefficient decreased a little.

NDVI, NDWI, and WDWI started with shallow values, increased and attained the highest values in September, and then decreased. The decrease rate for WDWI was fast rather than slow, as it was for NDVI, NDWI, and WDWI (Figure 5). S2REP behaved similarly to NDWI, except that the highest value was observed in July rather than September (Figure 5). The Sentinel-2 indices all followed the same trend with respect to various phenological stages, beginning with a low value in the germination stage, rising to the maximum value in the grand growth stage, and then decreasing as maturity approached. Saturations were found in all Sentinel-2 indices from tillering to grand growth, which occurred in July and September (Figure 5). In Figure 5, no significant difference was found due to the high growth of sugarcane biomass. The vegetation indices began to increase as the chlorophyll content increased during the tillering stage and began to fall as the plants matured and dried [80,81]. The NDWI and WDWI started rising as the proportion of water in the stem and leaf increased and the area per plant increased [82,83]. S2REP is affected by chlorophyll concentration, nitrogen, and growth status. The greater the S2REP value, the higher the chlorophyll concentration [42]. S2REP increased rapidly during the tillering stage as chlorophyll content in the stem and plant leaf increased.

Using machine learning approaches in this study, we investigated the potential of Sentinel-1, Sentinel-2, and combined data for assessing sugarcane phenology. Comparisons of accuracy and kappa values are given in Supplementary Figures S2 and S3. Table 2 and Figure 8 show the calculated accuracy, precision, specificity, sensitivity or recall, F score, area under curve (AUC) of receiver operating characteristic (ROC), and kappa value. Pairwise comparisons of accuracy and kappa value were carried out among the models as well as among the datasets using a re-sampling technique. The results are presented in Table 3.

Table 2. Performance metrics of classification models.

	Accuracy	Precision	Specificity	Sensitivity/ Recall	F1 Score	AUC	Kappa
Sentinel-1							
Decision tree	40.00%	29.30%	83.83%	40.83%	0.42	0.56	0.10
FRBS	56.00%	44.45%	87.77%	67.32%	0.48	0.59	0.28
Logistic	44.00%	35.04%	84.50%	33.33%	0.43	0.45	0.12
Naïve Bayes	52.00%	52.14%	86.25%	50.89%	0.48	0.58	0.34
Neural net	48.00%	29.87%	85.11%	23.89%	0.51	0.35	0.16
Random forest	60.00%	60.66%	88.50%	67.98%	0.60	0.83	0.44
SVM	52.00%	36.12%	86.19%	38.89%	0.46	0.47	0.23
Sentinel-2							
Decision tree	80.00%	93.78%	82.59%	80.06%	0.75	0.85	0.73
FRBS	64.00%	89.81%	63.57%	57.98%	0.54	0.69	0.52
Logistic	76.00%	92.99%	78.57%	71.53%	0.71	0.79	0.68
Naïve Bayes	56.00%	87.51%	59.20%	52.67%	0.44	0.66	0.42
Neural net	84.00%	95.35%	85.71%	82.76%	0.81	0.81	0.79
Random forest	72.00%	92.01%	75.45%	67.42%	0.66	0.73	0.63
SVM	76.00%	93.13%	79.02%	74.73%	0.71	0.75	0.68
Sentinel-1 and Sentinel-2							
Decision tree	76.00%	92.28%	69.05%	70.09%	0.68	0.83	0.63
FRBS	80.00%	93.31%	82.14%	79.49%	0.78	0.89	0.70
Logistic	80.00%	94.42%	80.95%	81.81%	0.78	0.83	0.76
Naïve Bayes	84.00%	96.11%	90.58%	87.21%	0.84	0.92	0.83
Neural net	88.00%	93.31%	77.38%	78.38%	0.75	0.87	0.70
Random forest	88.00%	95.74%	89.29%	88.32%	0.86	0.92	0.82
SVM	88.00%	95.74%	89.29%	88.32%	0.86	0.92	0.82

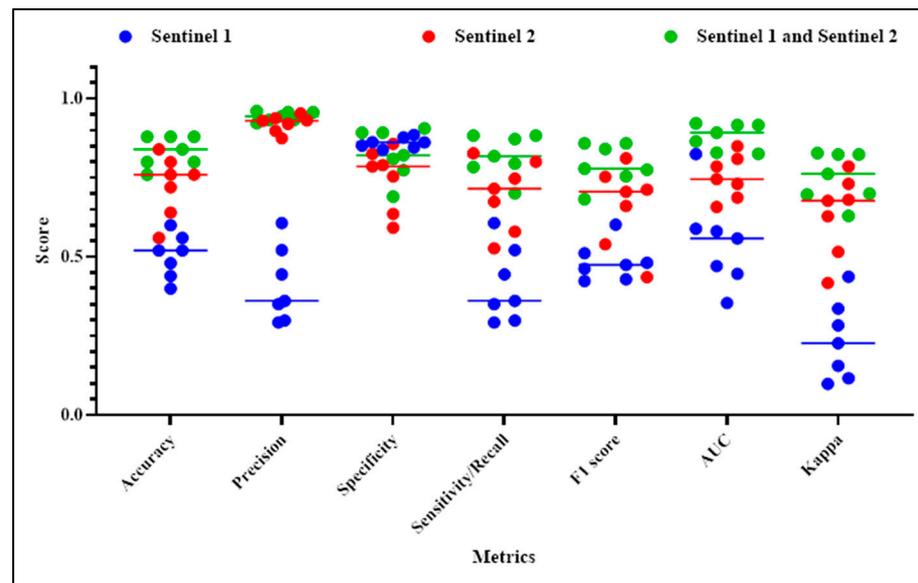


Figure 8. Data driven comparison of all metrics.

Table 3. Pair-wise test for accuracy and kappa value.

Models	Sentinel-1		Sentinel-2		Sentinel-1 and Sentinel-2	
	Accuracy	Kappa	Accuracy	Kappa	Accuracy	Kappa
Decision tree	f, c	f, c	b, a	b, a	d, d	d, b
FRBS	b, c	d, c	e, b	e, b	c, a	c, a
Logistic	e, c	e, c	c, b	c, b	c, a	b, a
Naïve Bayes	c, c	b, c	f, b	f, b	b, a	a, a
Neural net	d, c	e, c	a, b	a, b	a, a	c, a
Random forest	a, c	a, c	d, b	d, b	a, a	a, a
SVM	c, c	c, c	c, b	c, b	a, a	a, a

Note: All pairs were tested at 5% level of significance. The first letter stands for pairwise comparison of models for a dataset, while the second stands for pairwise comparison among datasets for a model. The presence of the same first letter in two models indicates that there is no difference in accuracy/kappa value between the pair of models in that particular dataset. The presence of the identical second letter in two datasets signifies that there is no difference in accuracy/kappa value between the pairs of data for that particular model. All of the pair comparisons are presented in Supplementary Tables S2–S20.

The results in Table 2 revealed that accuracy varied from 40% to 60%, 56% to 84% and 76% to 88% for Sentinel-1, Sentinel-2, and their combined data, respectively. It is clearly shown in Figure 8 that combining Sentinel-1 and Sentinel-2 data improved accuracy over using them individually. The model's role was less decisive when using both Sentinel-1 and Sentinel-2, but when a single data source was used, a difference in model performance was observed. Random forest was the best model for Sentinel-1 and combined data, whereas naive Bayes was the best model for Sentinel-2 data. For combined data, the neural net, random forest, and SVM models all had the same accuracy of 88%, which was the highest obtained value. Table 2 showed that the kappa value for Sentinel-1, Sentinel-2, and combined data ranged from 0.10 to 0.44, 0.42 to 0.79, and 0.63 to 0.83, respectively. The kappa value also indicated that combined Sentinel-1 and Sentinel-2 data were superior for sugarcane phenology prediction (Table 3). Based on the kappa value, random forest, neural net, and naive Bayes models outperformed the other models for Sentinel-1 data, Sentinel-2 data, and composite data, respectively (Tables 2 and 3). Table 3 clearly shows that except for the decision tree model, combined use of Sentinel-1 and Sentinel-2 data increased the accuracy and kappa value significantly.

Figure 8 depicts a data-driven comparison of all indices. Apart from precision and specificity, distinct differences could be detected for all metrics in respect to the dataset

used. The precision value for Sentinel-1 data was the lowest; however, the precision values for Sentinel-1 and combined data were inseparable. The specificity values for all three datasets were not distinguishable from each other. Sentinel-1 data had the lowest accuracy, recall, F1 score, AUC, and kappa values for phenology prediction, whereas combined data provided the highest values.

Based on the above reliable estimates, it can be inferred that composite data provide a more accurate picture of sugarcane phenology than individual data, and that Sentinel-2 indices are more informative than Sentinel-1 derived parameters.

3.2. Variable Importance

Variable importance specifies the number and combination of input features that are required to achieve satisfactory quality classification [84]. The mean rank of importance based on the mean decrease in the Gini index of the random forest model has been used to identify important features [85].

Using only Sentinel-1 parameters, the VV was the most important feature, followed by VH and VH/VV (Figure 9). For Sentinel-1 indices alone, NDVI was the most important feature, followed by NDWI. When using both Sentinel-1 and Sentinel-2, Sentinel-2 indices came out as highly important for identifying sugarcane phenological phases. NDVI appeared as the most important feature, followed by NDWI and VV. Figure 10 depicts mean rank importance of Sentinel features for each phenological phase. For Sentinel-1 parameters, VV was sensitive to germination, grand growth, and maturity stages, while VH was sensitive to tillering stages due to the double-bounce and volume scattering mechanism as the length of stem increased significantly [86]. Alpha was only sensitive to the maturity stage due to an increase in the total dominance scattering in this stage. For Sentinel-2 indices, NDVI was sensitive to the tillering and grand growth stages due to high chlorophyll content in stems and leaves. NDWI performed well in predicting germination and maturity stages due to better discrimination in water content than chlorophyll content. Based on this discussion, it is possible to conclude that NDVI, NDWI, VV, and VH were the most important features for capturing sugarcane phenology.

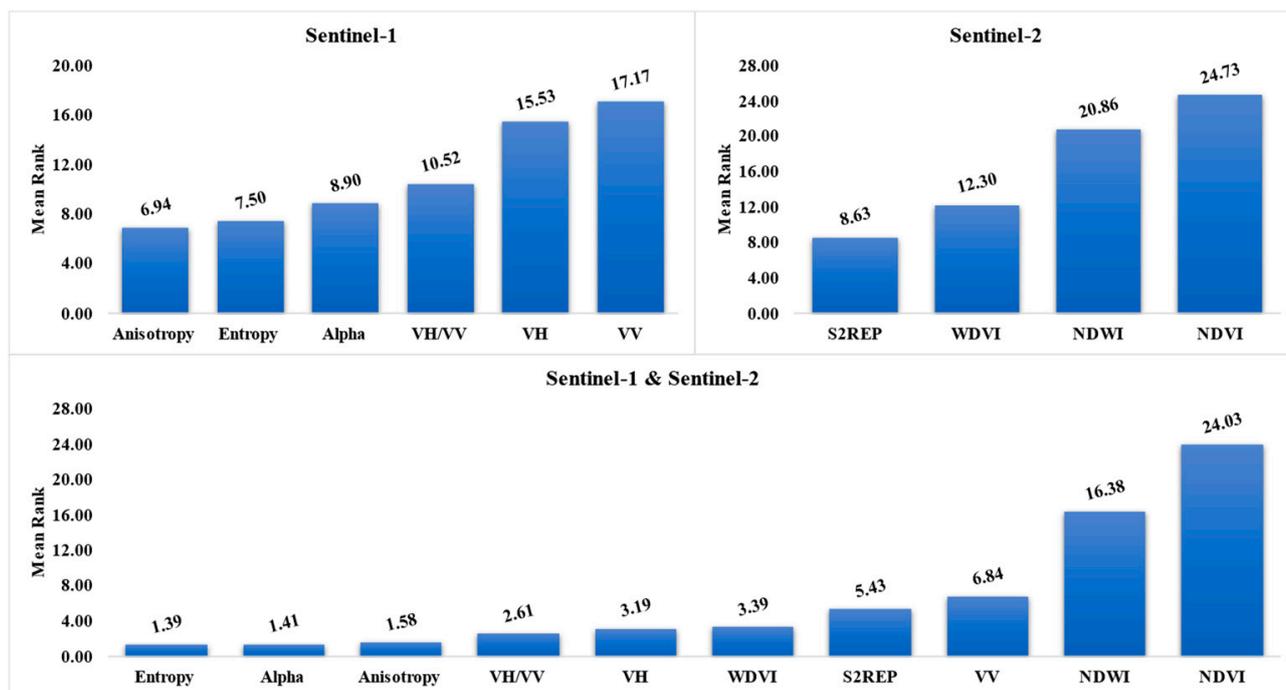


Figure 9. Mean rank of importance of different parameters and indices.

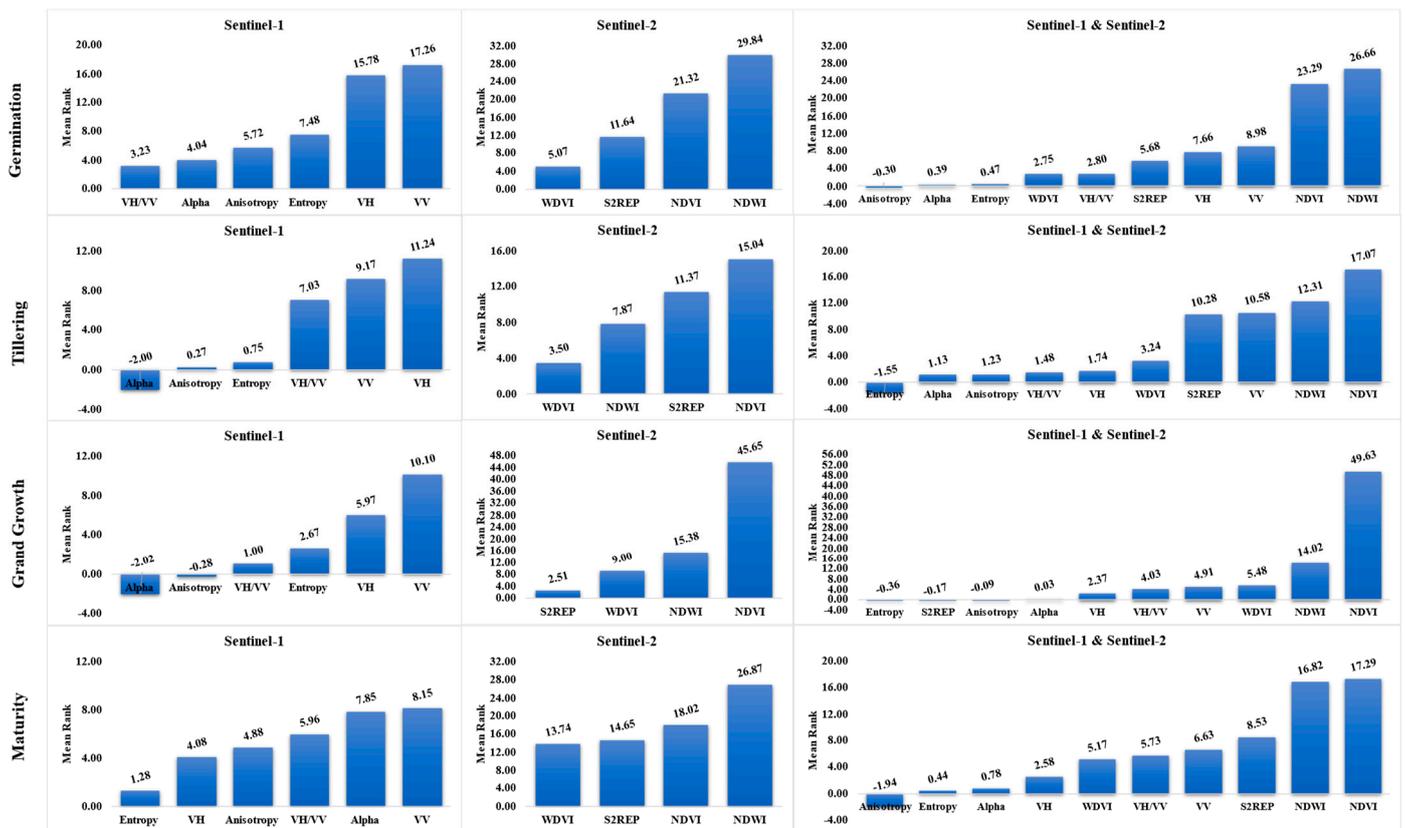


Figure 10. Mean rank of importance of different parameters and indices according to phenological phase.

4. Discussion

For predicting the phenological stages of sugarcane, combined use of Sentinel-1 and Sentinel-2 data outperformed the use of Sentinel-1 and Sentinel-2 data, as it took backscatter intensity, different scattering mechanisms as well vegetation vigour, chlorophyll content, and water content under consideration. The SAR parameters were mainly sensitive to the geometry and wetness of the crop and soil, whereas optical indices were affected by physiological parameters [21].

4.1. Sentinel-1 Based Parameters

Our research also found that the temporal behaviour of VV, VH, and VH/VV were similar, but that VV appeared to be best among the Sentinel-1 parameters to identify phenological stages. For alpha angle and entropy, similar temporal patterns were observed throughout the year. Anisotropy showed the opposite pattern with respect to alpha angle and entropy. Alpha and anisotropy were found to be more sensible to biomass growth. The Shannon entropy was found to be sensitive to low and medium crop biomass and useful in monitoring low biomass crops, but there was no significant variation response for advanced crop stages of wheat and mustard [25]. For mapping corn and soybean, Song et al. found that VV and VH polarizations were more important than the VH/VV ratio [24]. Tian et al. [87] and Gasparovic and Klobucar [88] discovered that VH and VV polarisation outperformed other Sentinel-1 parameters in rice and cotton vegetation studies, respectively. These findings are parallel with ours. Mercier et al. [21] discovered that the VH/VV ratio was most important for obtaining classifications of phenological stages for wheat and rapeseed crops, which contradicted our findings. Using seven machine learning models, Sentinel-1 based parameters produced 40.00–60.00%, 29.30–60.66%, 83.83–88.50%, 23.89–67.98%, 0.42–0.60, 0.35–0.83, and 0.10–0.44 values for accuracy, precision, specificity, recall, F1 value, AUC, and kappa value, respectively. The random forest model performed

the best, followed by the FRBS model, while the decision tree performed the worst among the evaluated models.

4.2. Sentinel-2 Based Indices

The temporal profiles of Sentinel-2 indices were consistent. Saturation was observed from tillering to grand growth. NDVI and WdVI were the most important features for Sentinel-2 indices. Chen et al. realised the potential of the red edge band and NDVI, and Hu et al. identified the superiority of the red edge band of Sentinel-2 images for rice and cotton field mapping, respectively [26,89]. The most important characteristics for wheat and rapeseed crop phenology were LAI, NDVI, and S2REP [21]. These findings contradicted our findings regarding the superiority of the red edge band/S2REP, but they supported the potential of NDVI for phenology mapping. Using Sentinel-1-based indices, machine learning models achieved accuracy, precision, specificity, recall, F1 value, AUC, and kappa values of 56.00–84.00%, 87.50–95.35%, 59.21–85.71%, 52.67–82.76%, 0.44–0.81, 0.66–0.81, and 0.42–0.79, respectively. The neural network model outperformed the other models, followed by the decision tree model, while the naive Bayes model was less effective than others.

4.3. Combined Sentinel-1 and Sentinel-2 Features

Song et al. [24] demonstrated that Sentinel-1 data alone can produce corn and soybean mapping with lower accuracy than optical data combining Landsat, Sentinel-2, Sentinel-1, and MODIS, which can achieve a potential accuracy of more than 95%. Mercier et al. [21] concluded that using Sentinel-1 and 2 data together was more accurate than using Sentinel-1 or Sentinel-2 data alone in identifying primary and secondary phenological stages in wheat and rapeseed. Chen et al. [26] observed the potential of Sentinel-2 images for rice phenology study, with overall accuracy and kappa coefficient values of 86.2% and 0.72, respectively. Hu et al. [89] used Sentinel-1 and Sentinel-2 data to study the rice phenology map and discovered overall accuracy and kappa coefficients of 0.932 and 0.813, respectively. Using Sentinel-2 based indices and Sentinel-1 based parameters, the accuracy, precision, specificity, recall, F1 value, AUC, and kappa value of the machine learning models increased to 76.00–88.00%, 92.28–96.11%, 69.50–90.58%, 70.09–88.32%, 0.68–0.86, 0.83–0.92 and 0.63–0.82, respectively. The random forest, neural network, and SVM models outperformed the other evaluated models for capturing sugarcane phenology. Complementary to our results, Feyisa et al. [90] inferred that SVM outperformed tree-based algorithms such as decision tree and random forest.

The majority of the literature on phenology studies reported that using both Sentinel-1 parameters and Sentinel-2 indices improved classification performance, and Sentinel-2 data were more useful than Sentinel-1 data. This study also supported the use of optical and SAR data in conjunction.

5. Conclusions

This study aimed to evaluate the potential of Sentinel-1 (GRD and SLC) data, Sentinel-2 (MSI level 2A) data, and their combined use to identify the phenological phases of sugarcane. Six Sentinel-1 indices with four Sentinel-2 indices were evaluated using seven machine learning algorithms. There was not much difference in the performance of the machine learning models. However, the random forest, neural net, and SVM models slightly outperformed the other evaluated models in terms of accuracy. The results show that the best Sentinel-2 based models were obtained using the NDVI and WdVI indices, and the most relevant features in Sentinel-1 based models were VV and VH polarization. The results obtained by combining Sentinel-1 and Sentinel-2 features were superior to those obtained by using Sentinel-1 and Sentinel-2 features separately.

This research can be extended by taking other relevant indices and parameters into account. These results were based on 40 sugarcane fields with an average size of 3–3.5 ha

and five satellite imagery datasets for different months. These results can be further validated through studies conducted over a larger area with a larger sample size.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs14143249/s1>.

Author Contributions: D.H., M.Y., S.K. and S.G. conceived the research. D.H. and M.Y. collected the data and designed the methodology. R.K.P. and M.Y. supported the empirical analysis. S.G. prepared the draft and edited the manuscript. S.K. supervised all activity. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are thankful to the Indian Institute of Remote Sensing (IIRS), ICAR-Indian Agricultural Statistics Research Institute and Haryana Space Application Centre (HARSAC).

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

Abbreviations

ANN	Artificial Neural Network
AUC	Area Under the ROC Curve
EVI	Enhanced Vegetation Index
FRBS	Fuzzy Rule Based Systems
GRD	Ground Range Detected
MODIS	Moderate Resolution Imaging Spectroradiometer
MSI	Multi-Spectral Instrument
NDVI	Normalized Difference Vegetation Index
NDWI	Weighted Difference Vegetation Index
RF	Random Forest
ROC	Receiver Operating Characteristic
RVI	Radar Vegetation Index
S2REP	Sentinel-2 Red-Edge Position Index
SAR	Synthetic Aperture Radar
SLC	Single Look Complex
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machine
SWIR	Short-Wave Infrared
VIRF	Visible Infrared Imaging Radiometer Suite
WDRVI	Wide Dynamic Range Vegetation Index

References

1. FAOSTAT. Available online: <https://www.fao.org/faostat/en/#home> (accessed on 26 May 2022).
2. Solomon, S. Sugarcane Agriculture and Sugar Industry in India: At a Glance. *Sugar Tech* **2014**, *16*, 113–124. [[CrossRef](#)]
3. Jyothi, K.C. Impact of Policy of Government on Import and Export of Sugar from India. *IOSR J. Econ. Financ.* **2014**, *3*, 40–42.
4. Lieth, H. *Phenology and Seasonality Modeling*; Springer Science & Business Media: Berlin, Germany, 2013; Volume 8, ISBN 364251863X.
5. Auffhammer, M.; Ramanathan, V.; Vincent, J.R. Climate Change, the Monsoon, and Rice Yield in India. *Clim. Chang.* **2012**, *111*, 411–424. [[CrossRef](#)]
6. Harvey, C.A.; Rakotobe, Z.L.; Rao, N.S.; Dave, R.; Razafimahatratra, H.; Rabarijohn, R.H.; Rajaofara, H.; MacKinnon, J.L. Extreme Vulnerability of Smallholder Farmers to Agricultural Risks and Climate Change in Madagascar. *Philos. Trans. R. Soc. B Biol. Sci.* **2014**, *369*, 20130089. [[CrossRef](#)] [[PubMed](#)]
7. Samui, R.P.; John, G.; Kulkarni, M.B. Impact of Weather on Yield of Sugarcane at Different Growth Stages. *J. Agric. Phys.* **2003**, *3*, 119–125.
8. Mall, R.K.; Sonkar, G.; Bhatt, D.; Sharma, N.K.; Baxla, A.K.; Singh, K.K. Managing Impact of Extreme Weather Events in Sugarcane in Different Agro-Climatic Zones of Uttar Pradesh. *Mausam* **2016**, *67*, 233–250. [[CrossRef](#)]
9. Diao, C. Remote Sensing Phenological Monitoring Framework to Characterize Corn and Soybean Physiological Growing Stages. *Remote Sens. Environ.* **2020**, *248*, 111960. [[CrossRef](#)]

10. Palaniswami, C.; Gopalasundaram, P.; Bhaskaran, A. Application of GPS and GIS in Sugarcane Agriculture. *Sugar Tech* **2011**, *13*, 360–365. [[CrossRef](#)]
11. Shihua, L.; Jingtao, X.; Ping, N.; Jing, Z.; Hongshu, W.; Jingxian, W. Monitoring Paddy Rice Phenology Using Time Series MODIS Data over Jiangxi Province, China. *Int. J. Agric. Biol. Eng.* **2014**, *7*, 28–36.
12. Wei, W.; Wu, W.; Li, Z.; Yang, P.; Zhou, Q. Selecting the Optimal NDVI Time-Series Reconstruction Technique for Crop Phenology Detection. *Intell. Autom. Soft Comput.* **2016**, *22*, 237–247. [[CrossRef](#)]
13. Liu, C.; Shang, J.; Vachon, P.W.; McNairn, H. Multiyear Crop Monitoring Using Polarimetric RADARSAT-2 Data. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 2227–2240. [[CrossRef](#)]
14. Ghaderpour, E.; Vujadinovic, T. The Potential of the Least-Squares Spectral and Cross-Wavelet Analyses for Near-Real-Time Disturbance Detection within Unequally Spaced Satellite Image Time Series. *Remote Sens.* **2020**, *12*, 2446. [[CrossRef](#)]
15. Ghaderpour, E. JUST: MATLAB and Python Software for Change Detection and Time Series Analysis. *GPS Solut.* **2021**, *25*, 85. [[CrossRef](#)]
16. Magdalena, L. Fuzzy Rule-Based Systems. In *Springer Handbook of Computational Intelligence*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 203–218. [[CrossRef](#)]
17. Sakamoto, T. Refined Shape Model Fitting Methods for Detecting Various Types of Phenological Information on Major US Crops. *ISPRS J. Photogramm. Remote Sens.* **2018**, *138*, 176–192. [[CrossRef](#)]
18. Sakamoto, T.; Wardlow, B.D.; Gitelson, A.A.; Verma, S.B.; Suyker, A.E.; Arkebauer, T.J. A Two-Step Filtering Approach for Detecting Maize and Soybean Phenology with Time-Series MODIS Data. *Remote Sens. Environ.* **2010**, *114*, 2146–2159. [[CrossRef](#)]
19. Standardi, L.; Karlsen, S.R.; Niedrist, G.; Gerdol, R.; Zebisch, M.; Rossi, M.; Notarnicola, C. Exploiting Time Series of Sentinel-1 and Sentinel-2 Imagery to Detect Meadow Phenology in Mountain Regions. *Remote Sens.* **2019**, *11*, 542. [[CrossRef](#)]
20. Song, Y.; Wang, J. Mapping Winter Wheat Planting Area and Monitoring Its Phenology Using Sentinel-1 Backscatter Time Series. *Remote Sens.* **2019**, *11*, 449. [[CrossRef](#)]
21. Mercier, A.; Betbeder, J.; Baudry, J.; Denize, J.; Leroux, V.; Roger, J.-L.; Spicher, F.; Hubert-Moy, L. Evaluation of Sentinel-1 and-2 Time Series to Derive Crop Phenology and Biomass of Wheat and Rapeseed: Northern France and Brittany Case Studies. In *Proceedings of the Remote Sensing for Agriculture, Ecosystems, and Hydrology XXI*, Strasbourg, France, 9–11 September 2019; Volume 11149, p. 1114903.
22. Gaetano, R.; Cozzolino, D.; D’Amiano, L.; Verdoliva, L.; Poggi, G. Fusion of SAR-Optical Data for Land Cover Monitoring. In *Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Fort Worth, TX, USA, 23–28 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 5470–5473.
23. Li, X.; Du, Z.; Huang, Y.; Tan, Z. A Deep Translation (GAN) Based Change Detection Network for Optical and SAR Remote Sensing Images. *ISPRS J. Photogramm. Remote Sens.* **2021**, *179*, 14–34. [[CrossRef](#)]
24. Song, X.-P.; Huang, W.; Hansen, M.C.; Potapov, P. An Evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS Data for Crop Type Mapping. *Sci. Remote Sens.* **2021**, *3*, 100018. [[CrossRef](#)]
25. Haldar, D.; Verma, A.; Kumar, S.; Chauhan, P. Estimation of Mustard and Wheat Phenology Using Multi-Date Shannon Entropy and Radar Vegetation Index from Polarimetric Sentinel-1. *Geocarto Int.* **2021**, 1–28. [[CrossRef](#)]
26. Chen, C.F.; Son, N.T.; Chen, C.R.; Chang, L.Y.; Chiang, S.H. Rice Crop Mapping Using Sentinel-1A Phenological Metrics. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*. [[CrossRef](#)]
27. Narin, O.G.; Abdikan, S. Monitoring of Phenological Stage and Yield Estimation of Sunflower Plant Using Sentinel-2 Satellite Images. *Geocarto Int.* **2020**, *37*, 1–15. [[CrossRef](#)]
28. Haldar, D.; Tripathy, R.; Dave, V.; Dave, R.; Bhattacharya, B.K.; Misra, A. Monitoring Cotton Crop Condition through Synergy of Optical and Radar Remote Sensing. *Geocarto Int.* **2022**, *37*, 377–395. [[CrossRef](#)]
29. Singh, D.; Singh, S.; Shekhar, C.; Singh, R.; Rao, V.U.M. *Agroclimatic Features of Hisar Region*; AICRP on Agrometeorology, Department of Agril Meteorology, College of of Agriculture, CCS Haryana Agricultural University: Haryana, India, 2010.
30. Ahlawat, I.; Sheoran, H.S.; Dahiya, G.; Sihag, P. Analysis of Sentinel-1 Data for Regional Crop Classification: A Multi-Data Approach for Rabi Crops of District Hisar (Haryana). *J. Appl. Nat. Sci.* **2020**, *12*, 165–170. [[CrossRef](#)]
31. Lee, J.-S.; Grunes, M.R.; de Grandi, G. Polarimetric SAR Speckle Filtering and Its Implication for Classification. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 2363–2373.
32. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, 1–33. [[CrossRef](#)]
33. Denize, J.; Hubert-Moy, L.; Betbeder, J.; Corgne, S.; Baudry, J.; Pottier, E. Evaluation of Using Sentinel-1 and-2 Time-Series to Identify Winter Land Use in Agricultural Landscapes. *Remote Sens.* **2018**, *11*, 37. [[CrossRef](#)]
34. Gascon, F.; Bouzinac, C.; Thépaut, O.; Jung, M.; Francesconi, B.; Louis, J.; Lonjou, V.; Lafrance, B.; Massera, S.; Gaudel-Vacaresse, A. Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sens.* **2017**, *9*, 584. [[CrossRef](#)]
35. Gamon, J.A.; Field, C.B.; Goulden, M.L.; Griffin, K.L.; Hartley, A.E.; Joel, G.; Penuelas, J.; Valentini, R. Relationships between NDVI, Canopy Structure, and Photosynthesis in Three Californian Vegetation Types. *Ecol. Appl.* **1995**, *5*, 28–41. [[CrossRef](#)]
36. Grace, J.; Nichol, C.; Disney, M.; Lewis, P.; Quaife, T.; Bowyer, P. Can We Measure Terrestrial Photosynthesis from Space Directly, Using Spectral Reflectance and Fluorescence? *Glob. Chang. Biol.* **2007**, *13*, 1484–1497. [[CrossRef](#)]
37. Karnieli, A.; Agam, N.; Pinker, R.T.; Anderson, M.; Imhoff, M.L.; Gutman, G.G.; Panov, N.; Goldberg, A. Use of NDVI and Land Surface Temperature for Drought Assessment: Merits and Limitations. *J. Clim.* **2010**, *23*, 618–633. [[CrossRef](#)]

38. Gao, B.-C. NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [[CrossRef](#)]
39. 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](#)]
40. Serrano, J.; Shahidian, S.; Marques da Silva, J. Evaluation of Normalized Difference Water Index as a Tool for Monitoring Pasture Seasonal and Inter-Annual Variability in a Mediterranean Agro-Silvo-Pastoral System. *Water* **2019**, *11*, 62. [[CrossRef](#)]
41. Bouman, B.A.M.; van Kasteren, H.W.J.; Uenk, D. Standard Relations to Estimate Ground Cover and LAI of Agricultural Crops from Reflectance Measurements. *Eur. J. Agron.* **1992**, *1*, 249–262. [[CrossRef](#)]
42. Guyot, G.; Baret, F. Utilisation de La Haute Resolution Spectrale Pour Suivre l'état Des Couverts Vegetaux. In Proceedings of the Spectral Signatures of Objects in Remote Sensing, Aussois, France, 18–22 January 1988; Volume 287, p. 279.
43. Clevers, J.; de Jong, S.M.; Epema, G.F.; Addink, E.A.; van der Meer, F.; Skidmore, A.K. Meris and the Red-Edge Index. In Proceedings of the Second EARSeL Workshop on Imaging Spectroscopy, Enschede, The Netherlands, 11–13 July 2000.
44. Rouse, B.T.; Wells, R.J.H.; Warner, N.L. Proportion of T and B Lymphocytes in Lesions of Marek's Disease: Theoretical Implications for Pathogenesis. *J. Immunol.* **1973**, *110*, 534–539.
45. Gao, B.C.; Goetz, A.F. Retrieval of equivalent water thickness and information related to biochemical components of vegetation canopies from AVIRIS data. *Remote Sens. Environ.* **1995**, *52*, 155–162. [[CrossRef](#)]
46. Clevers, J. The Derivation of a Simplified Reflectance Model for the Estimation of Leaf Area Index. *Remote Sens. Environ.* **1988**, *25*, 53–69. [[CrossRef](#)]
47. Maxwell, A.E.; Warner, T.A.; Guillén, L.A. Accuracy Assessment in Convolutional Neural Network-Based Deep Learning Remote Sensing Studies—Part 1: Literature Review. *Remote Sens.* **2021**, *13*, 2450. [[CrossRef](#)]
48. Sarlis, N.V.; Skordas, E.S.; Christopoulos, S.-R.G.; Varotsos, P.A. Natural Time Analysis: The Area under the Receiver Operating Characteristic Curve of the Order Parameter Fluctuations Minima Preceding Major Earthquakes. *Entropy* **2020**, *22*, 583. [[CrossRef](#)]
49. Cox, D.R. The Regression Analysis of Binary Sequences. *J. R. Stat. Soc. Ser. B Methodol.* **1958**, *20*, 215–232. [[CrossRef](#)]
50. Walker, S.H.; Duncan, D.B. Estimation of the Probability of an Event as a Function of Several Independent Variables. *Biometrika* **1967**, *54*, 167–179. [[CrossRef](#)] [[PubMed](#)]
51. Tabachnick, B.G.; Fidell, L.S.; Ullman, J.B. *Using Multivariate Statistics*; Pearson: Boston, MA, USA, 2007; Volume 5.
52. Good, I.J. Probability and the Weighing of Evidence. *Biometrika* **1951**, *38*, 485.
53. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin, Germany, 1999; ISBN 0387987800.
54. Nizar, A.H.; Dong, Z.Y.; Wang, Y. Power Utility Nontechnical Loss Analysis with Extreme Learning Machine Method. *IEEE Trans. Power Syst.* **2008**, *23*, 946–955. [[CrossRef](#)]
55. Berwick, R. An Idiot's Guide to Support Vector Machines (SVMs). Retrieved Oct. **2003**, *21*, 2011.
56. Hastie, T.; Tibshirani, R.; Friedman, J. High-Dimensional Problems: P n. In *The Elements of Statistical Learning*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 649–698.
57. Murthy, S.K. Automatic Construction of Decision Trees from Data: A Multi-Disciplinary Survey. *Data Min. Knowl. Discov.* **1998**, *2*, 345–389. [[CrossRef](#)]
58. Kotsiantis, S.B.; Zaharakis, I.; Pintelas, P. Supervised Machine Learning: A Review of Classification Techniques. *Emerg. Artif. Intell. Appl. Comput. Eng.* **2007**, *160*, 3–24.
59. Hunt, E.B.; Marin, J.; Stone, P. *Experiments in Induction*; Academic Press: New York, NY, USA, 1966.
60. Breiman, L.; Iha, R. *Nonlinear Discriminant Analysis via Scaling and ACE*; Department of Statistics, University of California: Los Angeles, CA, USA, 1984.
61. Liaw, A.; Wiener, M. Classification and Regression by RandomForest. *R News* **2002**, *2*, 18–22.
62. Probst, P.; Wright, M.N.; Boulesteix, A. Hyperparameters and Tuning Strategies for Random Forest. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2019**, *9*, e1301. [[CrossRef](#)]
63. Gupta, T.K.; Raza, K. Optimization of ANN Architecture: A Review on Nature-Inspired Techniques. In *Machine Learning in Bio-Signal Analysis and Diagnostic Imaging*; Academic Press: Cambridge, MA, USA, 2019; pp. 159–182.
64. Neocleous, C.; Schizas, C. Artificial Neural Network Learning: A Comparative Review. In Proceedings of the Hellenic Conference on Artificial Intelligence, Thessaloniki, Greece, 11–12 April 2002; Springer: Berlin/Heidelberg, Germany, 2002; pp. 300–313.
65. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
66. Zadeh, L.A.; Klir, G.J.; Yuan, B. *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers*; World Scientific: Singapore, 1996; Volume 6, ISBN 9810224214.
67. Sugeno, M.; Yasukawa, T. A Fuzzy-Logic-Based Approach to Qualitative Modeling. *IEEE Trans. Fuzzy Syst.* **1993**, *1*, 7–31. [[CrossRef](#)]
68. Pedrycz, W. *Fuzzy Modelling: Paradigms and Practice*; Springer Science & Business Media: Berlin, Germany, 1996; ISBN 0792397037.
69. Ishibuchi, H.; Nakashima, T. Effect of Rule Weights in Fuzzy Rule-Based Classification Systems. *IEEE Trans. Fuzzy Syst.* **2001**, *9*, 506–515. [[CrossRef](#)]
70. Lopez-Sanchez, J.M.; Vicente-Guijalba, F.; Ballester-Berman, J.D.; Cloude, S.R. Polarimetric Response of Rice Fields at C-Band: Analysis and Phenology Retrieval. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 2977–2993. [[CrossRef](#)]

71. Dey, S.; Bhogapurapu, N.; Bhattacharya, A.; Mandal, D.; Lopez-Sanchez, J.M.; McNairn, H.; Frery, A.C. Rice Phenology Mapping Using Novel Target Characterization Parameters from Polarimetric SAR Data. *Int. J. Remote Sens.* **2021**, *42*, 5515–5539. [[CrossRef](#)]
72. Varghese, A.O.; Joshi, A.K. Polarimetric Classification of C-Band SAR Data for Forest Density Characterization. *Curr. Sci.* **2015**, *108*, 100–106.
73. Mandal, D.; Kumar, V.; Ratha, D.; Dey, S.; Bhattacharya, A.; Lopez-Sanchez, J.M.; McNairn, H.; Rao, Y.S. Dual Polarimetric Radar Vegetation Index for Crop Growth Monitoring Using Sentinel-1 SAR Data. *Remote Sens. Environ.* **2020**, *247*, 111954. [[CrossRef](#)]
74. Harfenmeister, K.; Spengler, D.; Weltzien, C. Analyzing Temporal and Spatial Characteristics of Crop Parameters Using Sentinel-1 Backscatter Data. *Remote Sens.* **2019**, *11*, 1569. [[CrossRef](#)]
75. Fieuzal, R.; Baup, F.; Marais-Sicre, C. Monitoring Wheat and Rapeseed by Using Synchronous Optical and Radar Satellite Data—From Temporal Signatures to Crop Parameters Estimation. *Adv. Remote Sens.* **2013**, *2*, 33222. [[CrossRef](#)]
76. Cookmartin, G.; Saich, P.; Quegan, S.; Cordey, R.; Burgess-Allen, P.; Sowter, A. Modeling Microwave Interactions with Crops and Comparison with ERS-2 SAR Observations. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 658–670. [[CrossRef](#)]
77. Khabbazan, S.; Vermunt, P.; Steele-Dunne, S.; Ratering Arntz, L.; Marinetti, C.; van der Valk, D.; Iannini, L.; Molijn, R.; Westerdijk, K.; van der Sande, C. Crop Monitoring Using Sentinel-1 Data: A Case Study from The Netherlands. *Remote Sens.* **2019**, *11*, 1887. [[CrossRef](#)]
78. Wiseman, G.; McNairn, H.; Homayouni, S.; Shang, J. RADARSAT-2 Polarimetric SAR Response to Crop Biomass for Agricultural Production Monitoring. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4461–4471. [[CrossRef](#)]
79. Moran, M.S.; Alonso, L.; Moreno, J.F.; Mateo, M.P.C.; de La Cruz, D.F.; Montoro, A. A RADARSAT-2 Quad-Polarized Time Series for Monitoring Crop and Soil Conditions in Barrax, Spain. *IEEE Trans. Geosci. Remote Sens.* **2011**, *50*, 1057–1070. [[CrossRef](#)]
80. Ryu, J.-H.; Jeong, H.; Cho, J. Performances of Vegetation Indices on Paddy Rice at Elevated Air Temperature, Heat Stress, and Herbicide Damage. *Remote Sens.* **2020**, *12*, 2654. [[CrossRef](#)]
81. Gnyp, M.L.; Miao, Y.; Yuan, F.; Ustin, S.L.; Yu, K.; Yao, Y.; Huang, S.; Bareth, G. Hyperspectral Canopy Sensing of Paddy Rice Aboveground Biomass at Different Growth Stages. *Field Crops Res.* **2014**, *155*, 42–55. [[CrossRef](#)]
82. Mourad, R.; Jaafar, H.; Anderson, M.; Gao, F. Assessment of Leaf Area Index Models Using Harmonized Landsat and Sentinel-2 Surface Reflectance Data over a Semi-Arid Irrigated Landscape. *Remote Sens.* **2020**, *12*, 3121. [[CrossRef](#)]
83. Huang, J. *Vegetation Properties Relationships from Spectral Bands and Vegetation Indices from Operational Satellites*; The University of Manchester: Manchester, UK, 2006; ISBN 1392123283.
84. Kotsianti, S.B.; Kanellopoulos, D. Combining Bagging, Boosting and Dagging for Classification Problems. In Proceedings of the International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, Vietri sul Mare, Italy, September 12–14, 2007; Springer: Berlin, Germany, 2007; pp. 493–500.
85. Verikas, A.; Gelzinis, A.; Bacauskiene, M. Mining Data with Random Forests: A Survey and Results of New Tests. *Pattern Recognit.* **2011**, *44*, 330–349. [[CrossRef](#)]
86. Veloso, A.; Mermoz, S.; Bouvet, A.; le Toan, T.; Planells, M.; Dejoux, J.-F.; Ceschia, E. Understanding the Temporal Behavior of Crops Using Sentinel-1 and Sentinel-2-like Data for Agricultural Applications. *Remote Sens Environ.* **2017**, *199*, 415–426. [[CrossRef](#)]
87. Tian, H.; Wu, M.; Wang, L.; Niu, Z. Mapping Early, Middle and Late Rice Extent Using Sentinel-1A and Landsat-8 Data in the Poyang Lake Plain, China. *Sensors* **2018**, *18*, 185. [[CrossRef](#)]
88. Gašparović, M.; Dobrinić, D. Comparative Assessment of Machine Learning Methods for Urban Vegetation Mapping Using Multitemporal Sentinel-1 Imagery. *Remote Sens.* **2020**, *12*, 1952. [[CrossRef](#)]
89. Hu, Y.; Zeng, H.; Tian, F.; Zhang, M.; Wu, B.; Gilliams, S.; Li, S.; Li, Y.; Lu, Y.; Yang, H. An Interannual Transfer Learning Approach for Crop Classification in the Hetao Irrigation District, China. *Remote Sens.* **2022**, *14*, 1208. [[CrossRef](#)]
90. Feyisa, G.L.; Palao, L.K.; Nelson, A.; Gumma, M.K.; Paliwal, A.; Win, K.T.; Nge, K.H.; Johnson, D.E. Characterizing and Mapping Cropping Patterns in a Complex Agro-Ecosystem: An Iterative Participatory Mapping Procedure Using Machine Learning Algorithms and MODIS Vegetation Indices. *Comput. Electron. Agric.* **2020**, *175*, 105595. [[CrossRef](#)]