Categorisation of Vegetation Using Machine Learning and Remote Sensing Methods

Mansi Kambli * and Bhakti Palkar

Computer Department, K. J. Somaiya College of Engineering, Somaiya Vidyavihar University, Mumbai, Maharashtra, India
Email: mansi.mk@somaiya.edu (M.K.); bhaktiraul@somaiya.edu (B.P.)
*Corresponding author

***Abstract*—Farming is the main source of income, so crop development must be continuously watched and given the highest attention possible for the farmers. Precision agriculture is a farming management approach that uses multispectral satellite data to monitor, measure, and adjust to temporal and geographical variability in order to improve the sustainability of agricultural output. Sugarcane is a cash crop and is used in this study as researchers are focusing more on success in sugarcane development. Detecting dense and sparse vegetation for the individual plots of the farmers helps to understand that the plot area has favorable soil and water conditions for sugarcane growth. The sparse vegetation indicates that the area has slopes and water is not retained creating problems in the growth of the sugarcane. The Remote Sensing spectral bands are used and the vegetation indices like Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Ratio Vegetation Index (RVI) are used in the sugarcane canopy. EVI works for dense canopy and RVI for sparse vegetation as shown by the research done in this paper. The Machine Learning (ML) further also helps to detect the sparse and dense vegetation and the accuracy of all the classifiers is compared for the same. The survey on different Machine Learning techniques applied to remotely sensed data of sugarcane crops is done in this research work. This study aims to monitor sugarcane crop health by detecting sparse and dense vegetation using vegetation indices, and it evaluates the performance of different ML classifiers for precision agriculture. The same plot of the farmer can be monitored each month to find the change detection and further, the cause of sparse vegetation in the particular plot can be diagnosed with the help of enhanced vegetation indices in future work. To locate healthy vegetation, RS Sensing uses the Normalised Difference Vegetation Index (NDVI) and it gets saturated at grand growth stages so the novel method is enhance vegetation indices and ratio vegetation index which can be used to monitor at grand growth stages along with ML models as shown in this research work.**

***Keywords*—remote sensing, Geographic Information System (GIS), agriculture, machine learning, Normalized Difference Vegetation Index (NDVI)**

I. Introduction

Technology development is required to keep up with the rising demand for agricultural products. The cropland suitability and physical environmental variables are investigated along with Geographic Information System (GIS) and other forms of Remote Sensing (RS). Geographic Information System (GIS) processes the information gathered for visualization, analysis, and reporting through maps and charts. Precision agriculture has adapted statistical analysis or digital models using information technology. The researchers utilized the Information systems to precisely forecast crop productivity and land use. Therefore, information from farm databases is extracted using modern IT, data science, RS, and statistical techniques. It encourages academics to look into potent techniques for gathering spectral and field agricultural data. Brazil and India are the top two producers of sugarcane worldwide. The agro-industry economy is built on sugarcane and its products, such as ethanol, biofuel, and refined sugar. Sugarcane is a cash crop and in addition, it produces electricity and other essentials for biomass. Sugarcane is grown using two different agricultural techniques: seedlings and ratoon sugarcane. Compared to plant sugarcane, ratoon sugarcane matures more quickly. The crop output depends on the crop's health status and its regular monitoring. The crop data can be extracted by fusing RS data with image processing, data mining, Global Positioning System (GPS), and Geographic Information System (GIS). Traditionally, crop yield estimation was done by officials using sampling and enumeration methods and was done after the harvest of the crops. RS can be used before the harvest of the crops to detect minor changes in the crops effectively.

The dual polarization characteristics of sugarcane crops are used in the phenology of the crop. The polarization characteristics are low for the initial phase and when it matures, they vary with the Leaf Area Index (LAI) increase. NDVI is a quantitative measurement analysis of the crop and is used for the prediction of crop yield. The growth stages of the sugarcane crop are analyzed using different vegetation indices and estimating the leaf area index from it. The use of spectral bands along with spatial information helps in monitoring

Manuscript received October 6, 2024; revised November 5, 2024; accepted December 2, 2024; published April 25, 2025.

crop yield. Multi-temporal datasets and RS methods together can be used to predict crop yield. Zoning maps can also be made for harvest-related data and information on maximum sucrose content.

The main work is to identify various ML methods along with quantitative measurement indices such as the NDVI, which is used to monitor vegetation. From the papers surveyed, it is observed that RS with Geographic Information System (GIS) is used to find the relationship of various vegetation indices in a sugarcane crop. RS techniques contribute to sugarcane crop identification and estimation of crop condition assessment and raw cane yield forecasting. The data gathered from the different remotely connected sensors can help researchers better understand how soil, weather, and crops interact.

Water shortage is one of the main agricultural restrictions, especially in arid areas. The dry soils hence have minimal soil organic carbon, a weak soil structure, little soil biodiversity, and a high wind-driven soil erosion rate. The crop growth will be limited by increased aridity. And the availability of vital plant nutrients, endangering important ecological processes and services. This will lead to sparse vegetation. As the sensors physically if deployed on the farms can be affected by dust and other atmospheric effects, the satellite imagery data with its quality indicators like vegetation indices will give better results remotely about the dense and sparse canopy if monitored for each month. The identification of the gaps in the date of plantation and emergence period can be found by monitoring the growth which is helpful for the farmer to predict the harvest in time. Also, the gap is the sugarcane which is grown in the second year from the date of plantation and it is known as ratoon sugarcane. Its growth is faster as compared to the plant cane sugarcane. The challenges are by using vegetation indices through satellite imagery data the farmer can find the dense vegetation and sparse vegetation. This helps to determine whether the plot area is good in terms of soil and water for the sugarcane to grow and the sparse vegetation will indicate the farm that has slopes and water is not retained and creates problems in the growth of the sugarcane. For sparse vegetation, the Ratio vegetation indices will serve the purpose whereas for dense canopy the enhanced vegetation indices will give better results. The best month for harvest can be also found and ready to cut sugarcane from vegetation indices. The ground truth data is also not easily available in India and that is also one of the challenges which can be solved by satellite imagery data processing. The best quality indicator of satellite imagery data also needs to be found for monitoring the growth of sugarcane. When ML models like Random Forest, Support Vector Machine, and K Nearest Neighbour (KNN) are applied to RS vegetation indices (NDVI, EVI, and RVI), the accuracy and dependability of identifying changes in the health of sugarcane crops can be greatly improved. The NDVI, EVI and RVI are obtained by RS methods and are the feature extractors. The ML techniques are applied on the vegetation indices and the accuracy, precision and recall is assessed. The hybrid method of using RS and ML method in this research to detect dense and sparse vegetation is key factor of improving the crop management. Also Sustainable agriculture practices will be promoted and crop management decision-making will be enhanced by this method. The results of the work done are also added in the study, which shows the ML approaches applied to Sentinel-2 Earth Explorer and Sentinel-2 Multispectral Copernicus imagery data. The imagery data with cloud cover removed gave a good visualization of classes such as vegetation, barren, and water. In the remaining section, Section II illustrates a discussion of the review of the literature. Section III explains the methodology. Section IV is the results, whereas Section V explains the conclusion.

II. Literature Review

This section depicts the comparison between more conventional ML techniques and those used in deep learning when applied to remotely sensed data on sugarcane crops as well as other crops in agriculture and retrieves the evaluation parameters like features extracted, accuracy, which classifiers used, training samples, input imagery data; etc. The literature about the spatial, spectral, and temporal resolution of the crop is studied, and vegetation indices related to crop monitoring are assessed in the study.

A. RS and Machine Learning/DL Algorithms Applied on Crops

Singla *et al.* [1] proposed sugarcane harvest and crop development phases following atmospheric adjustments using the Support Vector Regression model. Wellens *et al.* [2] reported initial findings on how Sentinel-1 phenological indications might be incorporated into water crop models and illustrated Aquacrop models and the effect of water scarcity on agricultural production as well as mimicked irrigation’s impact on harvest size and quality. Yuan *et al.* [3] focused on identifying sugarcane in Southern China, with the Fusui region; serving as the research site. Sentinel-1 data were used to classify sugarcane samples, phenology, and polarizations. Ji *et al.* [4] compared the CASA and WOFOST models on wheat yield using Sentinel data. Lin *et al.* [5] found that sugarcane leaf area index and plant size are related using Envisat ASAR data. Martinez-Ferrer *et al.* [6] proposed multimodal satellite and meteorological data to estimate crop growth using the Gaussian process. The Gaussian process and Interpretive ML find the usage of each variable required for the prediction of yield as deep closed depressions or cut and fill in the fields cause a decrease in the yield as illustrated by Thomas *et al.* [7]. Virnodkar *et al.* [8] classified using deep learning models for different types of low-, medium-, and high-water stress sugarcane from satellite imagery. Virnodkar *et al.* [9] presented a review of the research on ML techniques to identify sugarcane crops by classifying RS images. Virnodkar *et al.* [10] has proposed CaneSat dataset and have used 2 dimensional Convolutional Neural Network (CNN) on RS satellite imagery of Sentinel-2 data. Acharya *et al.* [11] gives a summary of several current agricultural research projects that use GIS and remote sensing. Useya *et al.* [12] researched that for crop ensemble categorisation, decision-level data fusion using satellite images outperformed pixel-based data fusion. Li, and Chen [13] examined a number of indicators for remote sensing agricultural growth monitoring at various sizes. Khanal *et al.* [14] found an increasing trend in RS use over the past 20 years, with a significant rise in the use of unmanned aerial systems. Virnodkar *et al.* [15] showed that sugarcane crop fields may be successfully identified using the Sentinel-2 NDVI time-series with RF and SVM.Supavetch, Soravis [16] predicted in-season yields aids in creating seasonal business planning. The yield prediction before the harvest season was suggested by Mello *et al.* [17]. Virnodkar *et al.* [18] attempts to provide a comprehensive overview of the popular techniques for agricultural water stress monitoring that combine machine learning and remote sensing. Sahoo *et al.* [19] illustrated the application of Hyperspectral imagery data in agriculture. Ashcraft and Karra [20] presented a crop simulation environment using OpenAI Gym interface and deep reinforcement learning algorithms to optimize crop yield, demonstrating their potential to discover new strategies while minimizing constraints like water and fertilizer usage. Tang *et al.* [21] explored the concept of digital agriculture. Gopal [22] evaluates machine learning algorithms for accurate crop yield prediction in agriculture. Deng *et al.* [23] have used various numerical models such as decision trees and multiple linear regressions to keep an eye on the sugarcane plantation process area based on NDVI time series and agro-meteorological data. ValleGonçalves *et al.* [24] presented numerical models used to track sugarcane productivity using time series NDVI and agro meteorological data. Lu and Weng [25] also proposed different methods of ML for boosting the efficiency related to crops. Different RS applications in cultivating sugarcane were addressed by SomArd *et al.* [26]. Panwar *et al.* [27] proposed biological variables using a Gaussian process model based on ML in sugarcane crops and used a combination of vegetation indices as the parameter for it. Kai *et al.* [28] stated the multivariate methods for the identification and discrimination of sugarcane variety based on different vegetation indices. It also helps in harvesting the right variety of sugarcane based on soil and climate conditions. Further, using vegetative indices, an ideal model may be developed for determining when to harvest each sugarcane variety to maximize yield and sucrose content in the final product. This was proposed by Jamnani *et al.* [29]. Kamilaris and Prenafeta-Bold [30] and Victor *et al.* [31] have surveyed the application of deep learning to agricultural satellite images. Deep learning techniques have outperformed ML models and are efficiently used in agriculture. The growth stages of sugarcane were monitored using Lidar datasets and RS techniques, and further crop yield was found by Villareal *et al.* [32].

Sentinel-1 and Sentinel-2 datasets were combined to monitor the growth of sugarcane plantations and 7 ML methods were applied for the same by Yeasin *et al.* [33] Sentinel-1 and Sentinel-2 produce more precise results as suggested by Yeasin *et al.* [33]. Canata *et al.* [34] claimed that Sentinel-2's multitemporal imagery data was integrated with the random forest regression technique to create prediction yield models for industrial sugarcane farms. Scrivani *et al.* [35] illustrated the use of 3 variables and planted area for estimating sugarcane production and numerical models were generated by multiple linear regression analyses. Abdel-Rahman and Ahmed [36] stated the challenges and opportunities persistent to RS applications in sugarcane production. The temporal analysis of different crops like banana, cotton and sugarcane using RADARSAT-2 data is suggested by Kumar *et al.* [37].

Racine Ly *et al.* [38] illustrated ML along with RS for forecasting of sugarcane production, which will help decision-makers or policymakers to make decisions before the harvest period of the crop and not wait until the crop is ready to cut. The goal of Krupavathi *et al.* [39] is to use features derived from RS and previous crop yields and develop models to predict sugarcane crop yields.

A survey on research trends in smart agriculture is depicted in the study of Mitra *et al.* [40]. The challenges and problems related to smart agriculture are also addressed. The data from the spectral bands are mainly utilised to analyze maize crop, and then the Index of Leaf Area is calculated to find the crop growth.

LAI estimation for the development of the crop is done by using data from both the red and red-edge spectral bands to analyse maize crops and the study is proposed by Xie *et al.* [41] and Sharif [42]. Also, Sharif [42] showed that maize nitrogen uptake may be better measured using the spectral bands. Lozano-Garzon *et al.* [43] suggested providing a software application that uses RS and ML approaches to detect sugarcane crops CNN's deep learning technology that automatically detects palms and locates them is suggested by Ammar *et al.* [44].

The impact of climate change on sugarcane output is described by Singh *et al.* [45]. Mamatkulov *et al.* [46] monitored and forecasted crop yields in Jarkurgan district's high and low-productive cotton patches using cutting-edge technologies like GIS and Manjula *et al.* [47] illustrated the process of digitising a forest map and turning the georeferenced map into attribute data that can be utilised to create a spatial database and conduct spatial analysis.

Simoes *et al.* [48] proposed to monitor and distinguish crops using multitemporal Radar Sat fine mode data. The improved Change Detection technique for satellite image analysis based on Normalised Difference Vegetation Index (NDVI) is presented in this article by Gandhi *et al.* [49]. Sandeep *et al.* [50] suggested RS-Based prediction of sugarcane yield through a combination of ML Methods.

Ballo *et al.* [51] presents a hybrid approach for classifying mango leaf diseases that combines machine learning algorithms with deep learning feature extractors. In order to detect and classify mango leaf illnesses, study first employed three machine learning algorithms: Random Forest, SVM, and XGBoost. Next, they combined feature extractors from algorithms like VGG16, VGG19, and ResNet50V2 in an effort to maximize the outcomes. The findings indicate that there is a great deal of promise for identifying mango diseases using the hybrid approaches.

Bunyang *et al.* [52] investigates the efficacy of contrastive pre-training using SimCLR for self-supervised learning in the categorisation of plant disease images. It looks into label efficiency, fine-tuning on labelled samples and the findings demonstrate that self-supervised learning on plant disease can overcome the learning bottleneck caused by huge, expensive, and time-consuming photos and is on par with supervised training.

An evaluation by comparison of ML learning algorithms applied to remotely sense agricultural data is shown in Table I.

TABLE I. A COMPARATIVE ANALYSIS OF ML ALGORITHMS APPLIED TO REMOTELY SENSED AGRICULTURAL DATA

Sr No.	Authors, Year	Crop	ML models used	Dataset	Research findings	Accuracy/Inference
1.	Singla *et al.* [1]	Sugarcane	SVR model	Landsat imagery	GIS was used to correlate RS data with sugarcane Brix values.	90.4
2.	Wellens *et al.* [2]	Sugarcane	SAR time series analysis	Sentinel-1 and Landsat imagery	The absorption stages of the Aqua crop for phenological crops were determined using the SAR time series	83% planting 94% harvesting
3.	Yuan *et al.* [3]	Sugarcane, rice, banana	cross-polarized VH	Sentinel-1 data	The different polarized characteristics of the crops were analyzed	88.07%
4.	Ji *et al.* [4]	Wheat	CASA-WOFOST	Sentinel-2 data	A comparison of the model's CASA and WOFOST was done on wheat yield by assimilating Sentinel data time series	84%
5.	Lin *et al.* [5]	Sugarcane	HV/HH polarization	ENVISAT ASAR data	Envisat ASAR data is used for relating sugarcane leaf area index to plant size	78%
6.	Martinez-Ferrer *et al.* [6]	Corn, wheat	Gaussian process model	USDA-NASS	The use of the Gaussian process for estimating crop development using multisensory satellite and meteorological data	82%
7.	Jones *et al.* [7]	Cotton	Interpretive ML	Landsat-5 imagery	It aids agronomists in overcoming yield limits by implementing management strategies.	75%
8.	Virnodkar *et al.* [8]	Sugarcane	ML algorithms	Remote sense satellite imagery	Sugarcane field classification by deep learning use satellite imagery.	74.6 to 100
9.	Virnodkar *et al.* [9]	Sugarcane	SVM, Random Forest (RF)	Sentinel 2A	Phenology of the crop	SVM-81.86
10.	Virnodkar *et al.* [10]	Crops	ML	Sentinel data imagery	Analysis of ML and RS for Water Scarcity Management	83.30%
11.	Useya *et al.* [12]	Crops	Ensemble classifier	Landsat-8, Landsat-7, Sentinel-2 data	Decision-level data fusion of satellite imagery was better than pixel-based data fusion for crop ensemble classification.	Pixel level fusion −82.5 Decision-level fusion
12.	Supavetch [16]	Sugarcane	Linear model	High resolution satellite imagery	In-season yield prediction helps in making business plans for the season	low (8 tonnes) Excellent (more than 18 tonnes)
13.	Mello *et al.* [17]	Sugarcane	Linear model	MODIS (MOD13)	Yield estimates can be obtained before the harvest season	70% estimated yield
14.	Sahoo *et al.* [19]	Crops	Empirical models	Hyperspectral data imagery	Applications of hyper spectral RS in agriculture	96%
15.	Tang *et al.* [21]	Crops	Empirical models	Satellite data imagery	The concept of Digital Agriculture is explored	The full use of digital machinery
16.	Gopal [22]	Crops	ML algorithms	Meteorological data	The best feature subsets for predicting agricultural yield with ML methods are analyzed for their performance.	R value −93%
17.	ValleGonçalves *et al.* [24]	Sugarcane	Multiple Linear Regression	Agro Meteorological data	The study used time series NDVI and agro meteorological data to present numerical models applied to monitor sugarcane productivity.	90%
18.	Lu and Weng [25]	Crops	Different ML techniques	Remotely sensed data	Analysis of Image Classification Techniques to Boost Efficiency	88%
19.	Ekta Panwar *et al.* [27]	Sugarcane	Gaussian process regression	Sentinel-2 data	Monitoring of sugarcane crops based on vegetation indices and biochemical variables	R^2 = 0.7
20.	Rahimi *et al.* [29]	Sugarcane	Quadratic methods	Landsat-8 data	A zoning map showing when different locations will be ready to harvest to maximize sucrose content.	R^2 = 0.855
21.	Villareal *et al.* [32]	Sugarcane	SVM model	LIDAR datasets	The sugarcane mapping and classification using RS	98.4%
22.	Yeasin *et al.* [33]	Sugarcane	ML models	Fused data of	Seven ML techniques were implemented	88%

				Sentinel-1 along with Sentinel-2	and compared for the phenology assessment of sugarcane	
23.	Tatiana *et al.* [34]	Sugarcane	RF and multiple Linear Regression	Sentinel-2 data	The spectral bands show more contribution in predicting the yield of sugarcane crops compared to vegetation indices.	$R^2 = 0.70$
24.	Scrivani *et al.* [35]	Sugarcane	Linear Regression	MODIS data	The numerical models predicted sugarcane yield that use indices and planted area as independent variables.	$R^2 = 0.94$
25.	Abdel-Rahman *et al.* [36]	Sugarcane	Empirical Models	Satellite imagery from NOAA/AVHRR	Sugarcane yield forecasting and variety identification using RS are discussed.	The application of RS techniques to sugarcane
26.	Kumar *et al.*[37]	Crops	Eigenvector decomposition	RADARSAT-2 data	The temporal behavior of the crops with crop growth stages is monitored	80%
27.	Xie *et al.* [41]	Maize	Sensitivity analysis	Rapid Eye images	Red and red edge reflectance indices have been refined for use in estimating LAI and monitoring the stages of the crop.	NDVI = 0.7
28.	Sharifi *et al.* [42]	Maize	Predictive model	Sentinel-2 data	Vegetation indicators would be the most accurate indicators of maize nitrogen absorption when using red-edge and near-infrared bands.	$R^2 = 0.91$
29.	Lozano-Garzon *et al.* [43]	Sugarcane	Predictive models	Landsat-8, Sentinel-2 data	To propose software tool for identifying sugarcane crops through RS and ML techniques.	91%
30.	Singh *et al.* [45]	Sugarcane	No models	Meteorological data (IMD)	It is explained how the production of sugarcane is affected by climate change.	Rainfall and temperature are important parameters
31.	Simoes *et al.* [48]	Crops	Quantitative and qualitative analysis	Radar Sat SAR imagery data	For crop discrimination and monitoring using Radar Sat fine mode multitemporal data.	95%
32.	Sandeep *et al.* [50]	Crops	Random Forest	Landsat-8	RS-Based Sugarcane Yield Prediction Using an Ensemble of ML Techniques	$R^2 = 0.94$

B. *Observations of ML Techniques on Sugarcane Crop*

1. RS techniques are contributing to sugarcane crop identification and estimation of the crop condition assessment and raw cane yield forecasting. Further ML techniques can be explored more on the same applications.
2. In future research ML and RL techniques can differentiate between different types of sugarcane.
3. The growth stages of the sugarcane crop are analyzed by using different vegetation indices and estimating the Leaf area Index can be done further by using ML methods.
4. The zoning maps can be made for the harvest-related data and to get information on maximum sucrose content by using ML methods.
5. Decision-level data fusion of satellite imagery was found better than pixel-based data fusion for crop ensemble classification. The accuracy can be enhanced further by ML methods.
6. The RS-based Sugarcane Yield Prediction is done using an Ensemble of ML Techniques.
7. To propose software tool for identifying sugarcane crops through RS and ML techniques.
8. The numerical models predicted sugarcane yield that used NDV, WRSI, and planted area as independent variables.

C. *RS and Deep Learning Algorithms Applied on Crops*

An evaluation based on comparison of deep learning algorithms applied to remotely sense agricultural data is shown in Table II.

TABLE II. A COMPARATIVE ANALYSIS OF DEEP LEARNING (DL) ALGORITHMS APPLIED TO REMOTELY SENSE AGRICULTURAL DATA

Sr No.	**Authors, Year**	**Crop**	**Deep learning Models used**	**Dataset**	**Research findings**	**Accuracy/Inference**
1	Virnodkar *et al.* [8]	Sugarcane	Resnet, Densenet	Sentinel-2 imagery	Findings for different types of low, medium, and high stress sugarcane from the satellite	80.53
2	Virnodkar *et al.* [10]	Sugarcane	2D CNN	CANESAT	The classification of sugarcane and non-sugarcane crops using the CaneSat dataset.	88.46%
5	Ashcraft *et al.* [20]	crops	Deep Reinforcement algorithms	Greenhouse environment	To help crop yield optimization	A plant simulation model was made
4	Li and Chen [25]	wheat	Hyper spectrometer	Multispectral image sensing using an HJ-IA satellite	Wheat development is remotely tracked for stages at the canopy scale using many indicators.	SAVI (L = 0.3) for the elongation phase. SAVI (L = 0.2 for jointing
5	Kai *et al.* [28]	Sugarcane	Multivariate methods	Landsat-7 data	Discrimination of sugarcane variety by RS	95%
6	Kamilaris *et al.* [30]	Crops	Deep learning	Satellite data	The deep learning how it is	95% and above

			models	imagery	used in agriculture is explained	
7	Victor *et al.* [31]	Crops	Deep learning models	Sentinel, Landsat and MODIS	The study explores the DL techniques on images of the satellite in agriculture domain.	A review of Deep Learning in Agriculture
8	Ly *et al.* [38]	Crops	Artificial Neural Network	MOD13A2 dataset	The production of the crop is determined using RS and ML methods	RS and ML for decision making in agriculture
9	Krupavathi *et al.* [39]	Sugarcane	Artificial Neural Network	LANDSAT-8 imagery data	The goal is to use RS factors and previous yield data to create simplified yield prediction models for the sugarcane crop.	95%
10	Mitra *et al.* [40]	Crops	Deep Learning	Satellite imagery data	The survey on research trends in Smart Agriculture	Challenges in agriculture are discussed
11	Ammar *et al.* [44]	Palm tree	Deep learning methods	Aerial view dataset	Automatic recognition and geolocation of palm trees by DL methods	93%
12	Singh *et al.* [45]	Sugarcane	The analysis of meteorological data	Meteorological data (IMD)	The environmental impacts on the sugarcane production are discussed.	Meteorological data are obtained and the impact of climate is studied
13	Ballo *et al.* [51]	Mango	ML and DL models(Hybrdi model)	MangoLeafBD dataset	A major step toward enhancing plant disease prevention, crop preservation, and mango disease diagnosis and management.	XGBoost −90.71 VGG16-SVM −97.75 VGG19-SVM −97.12 ResNet50V2-SVM 97.31
14	Bunyang *et al.* [52]	Crops	Self-supervised (SimCLR)	Plant disease dataset (Kaggle)	Self-supervised learning's performance on plant disease was on par with the supervised training method.	Self-supervised Resnet50 −85% Self-supervised VGG16 −80%

D. Observations of Deep Learning Algorithms Applied on Sugarcane Crop

1. The DL models are used for the classification of sugarcane crops and can be.
2. explored more in terms of accuracy.
3. The discrimination of sugarcane variety can be done by RS and Deep learning techniques.
4. The Deep learning techniques can be applied for low, high, and medium water stress.

III. Materials and Methods

This section acts as a step-by-step guide on how we can use the publicly available satellite imagery data and different ML techniques applied to the vegetation indices. The vegetation index is a quantitative measurement of the crop, and the NDVI vegetation index is explored for different classes. The remotely sensed data from Earth Explorer and Copernicus are used in this implementation for the area of Belagavi in Karnataka with latitude 16.332 and longitude 75.2669. The satellite data from Earthexplorer and Copernicus from January 2022 are taken for the analysis of vegetation. The selection of ML algorithms, selection of methodologies for crop vegetation, and gathering of information from satellite data are all part of the methodology.

The RGB color space image downloaded from the satellite imagery, as shown in Fig. 1, is taken, pre-processing is done, and the cloud cover is removed. The classes that are not required for the analysis are removed and for each class that is selected training samples are made for evaluation. The input image is pre-processed and on the image with cloud cover removed further, the vegetation indices are evaluated for finding classes like water, barren land, and vegetation

The rain falls during certain months and cloud cover impacts satellite imaging data, which in turn affects NDVI values and produces inaccurate canopy characteristics of the sugarcane crop. As the sugarcane crop must be monitored for a whole year, accurate atmospheric correction must also be carried out.

A. Pre-processing

The Google Earth Engine JavaScript script carries out the following functions:

1. Establishing the ROI (region of interest):
   The region of interest is represented in acres by a polygon ("ROI"). The coordinates for the lower-left, lower-right, upper-right, and upper-left corners are used to build the polygon.
2. Establishing Time Frame:
   The time frame for which Sentinel-2 imagery is requested is defined by the parameters "start Date" and "end Date".
3. Sentinel-2 Image Filtering:
   The defined ROI and date range are used to filter Sentinel-2 images. ".filter Bounds (ROI)": Removes images that cross the specifiedROI. filter Date (start Date, end Date): Filters images that fall within time frame
4. Band Selection Function:
   A function to choose particular bands from the Sentinel-2 picture is defined via the "selectBands" function.

The "B8" (near-infrared) band and Red band ("B4") along with blue band ('B2') is chosen, A new picture with just the chosen bands is returned by the function.

5. Using the Collection to Apply Band Selection: "selected Bands Collection: Uses .map (select Bands)" to apply the "select Bands" method on the complete image collection.
6. Using Google Drive to Export Selected Bands: "Export.Image.toDrive": Exports a GeoTIFF file of the median picture with chosen bands to Google Drive.
7. Cloud filtering using the Median method: In Median method the neighboring pixel values are used and median is calculated for cloud filtering and filling the data with median values. Also it is better as compared to Histogram minimum method. Also, NDVI values are increasing after atmospheric correction using Median method.

A similar methodology is used for the image with cloud cover for a better assessment of the crop. Red-edge and near-infrared bands are used for analysis in Copernicus satellite imagery data. They provide better visualization compared to Earth Explorer satellite imagery data. The tabular data of the vegetation indices for sugarcane crop from Jan 2022 for sugarcane crop monitoring for 12 months is shown in Fig. 1 and the observations are shown in the Results.

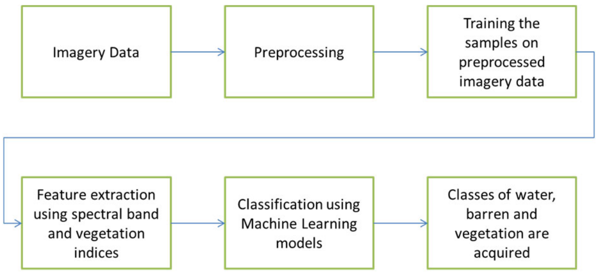

Fig. 1. Methodology for data acquisition and processing of satellite imagery data.

First, download the TIF image as shown in Fig. 2 from Earth Explorer [53] https://earthexplorer.usgs.gov/

Step 1: Enter latitude and longitude and make cloud cover 0, and the radius value is inserted. In addition, polygons with a circle as their center can have their radius and center point defined in decimal degrees of latitude and longitude.

Step 2: After that, select the dataset Sentinel and the acquisition date is selected

Step 3: Check the footprint of the region and download the TIF image.

B. Datasets

Copernicus, developed by the European Space Agency (ESA) [54], uses Sentinel-2 data to monitor Earth from space and can be downloaded for free at https://scihub.copernicus.eu/andhttps://earthexplorer.usgs.gov/ used in this paper.

The satellite's opto-electronic multispectral sensor allows for surveying with a Sentinel-2 resolution of 10 to 60 m in the visible, Near-Infrared (VNIR), and Short-Wave Infrared (SWIR) spectral zones. It has 13 spectral channels, which minimizes the impact on atmospheric photography quality while ensuring that variations in vegetation state, including temporal changes, are captured.

The tiff images were taken from the satellite imagery site and jpg images were taken from the farmers through GPS map camera and further they were mapped and georeferenced for the analysis purpose. The resolution was 10×10 pixels for tiff image.

The geographic coordinates (latitude and longitude) were taken from the farmer's. Then the farmer's jpg image was converted to a kml file using a python code to get the farm placeholder. Put the kml file path in the merged coordinate's python code to get the coordinates of the whole area. Then the datasets from the Sentinel 2 dataset based on the merged coordinates were downloaded for further analysis.

The sugarcane crop has to be monitored for 12 months to 15 months and satellite imagery data of Sentinel-2 was collected over multiple growing seasons for individual plot of the farmer. In this research it is monitored from Jan 2022 to March 2023. The month from germination stage to harvest stage is taken into account for NDVI, EVI and RVI calculation as shown in Tables III–V.

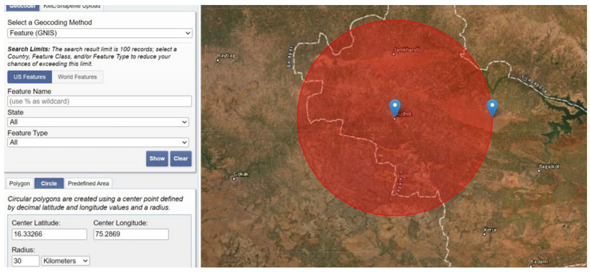

Fig. 2. TIF image downloaded from earth explorer [53].

C. ML Algorithms

The ML strategies for classifying sugarcane crops using remotely sensed data are illustrated in this study.

- **Support Vector Machine (SVM)**

SVM classification is a supervised learning model used for regression and classification. A statistical learning technique called support vector machines can more accurately classify heterogeneous data without making assumptions about the distribution of the input data. The goal of an SVM learner is to reach the Optimal Separation (OSH), which is a choice. The barrier between classes with a maximum margin reduces classification error during training and subsequently generalizes to new data. Support vectors are utilized to optimize the classifier's margin. Support vectors are data points that primarily aid in fitting the hyper plane and are located closer to the margin. Since they don't affect the hyper plane's position or orientation, other data points are eliminated. SVM classification results are better than random forest, maximum likelihood, K Nearest Neighbor, and Iterative Self Organising (ISO) clustering. The RGB color space image and the output give a proper visualization of water, barren land, and cultivated land after applying the SVM classifier

- **Random Trees**

The Random Trees classification method is a supervised ML classifier that builds numerous decision trees, assigns each tree a random subset of factors, and uses the most common tree output to determine the classification as a whole. The random tree classification gives better results than the maximum likelihood and will perform on the random selection of training pixels. The vegetation class is not visualized totally as compared to the SVM classifier. The water and barren land are classified to some extent.

- **Maximum Likelihood**

A parametric classifier known as the Maximum Probability Algorithm (MLA) assigns a pixel to a class according to the likelihood that it will belong to a class whose mean and covariance are shown to constitute a normal distribution in multispectral feature space. MLA is a simple classifier that was commonly used before invention. Under the premise that the statistics of each class in each band are regularly distributed, maximum likelihood classification can provide an approximate likelihood that a given pixel is a member of a certain class. Pixels are then placed in the most likely category.

- **K nearest Neighbour**

K nearest neighbor stores all the given input data and at the time of classification it performs the actions on the dataset. The K value can be 1, 2, 3 or any and depending on K's nearest neighbors the classification is done. The dataset if it is large then KNN would be more effective.

- **Iterative Self Organising**

ISO clustering is an unsupervised clustering technique. The vegetation results are obtained to some extent but not precise. The iterative self-organizing method of clustering is shortened to the iso prefix of the isodata clustering algorithm. Every iteration of this kind of clustering involves recalculating the means for each class and assigning all samples to pre-existing cluster centers ISO clustering is unsupervised and gives misclassified results. The image with cloud cover is interpreted as vegetation and more work has to be done on unsupervised classification techniques. It is an iterative self-organizing way of performing clustering

D. Performance Metrics

The performance is checked by using the evaluation parameters like Accuracy, Precision and recall which are briefly explained in this study.

Accuracy is a statistic commonly used in deep learning and ML models to evaluate the effectiveness of a categorization process.

$$Accuracy = \frac{(TN + TP)}{(TP + FP + TN + FN)} \tag{1}$$

A binary classification model's efficacy is evaluated using a measure called precision

It quantifies the percentage of actual positive cases among all predicted positive examples. Stated differently, it assesses the model's ability to correctly detect positive cases while lowering false positives.

$$Precision = \frac{TP}{(TP + FP)} \tag{2}$$

Recall is the model's ability to identify positive examples while lowering false negatives, which calculates the percentage of True Positives (TP) among positive and negative cases. This data is then used to evaluate the efficacy of models for binary classification

$$Recall = \frac{TP}{(TP + FN)} \tag{3}$$

E. Selection of Vegetation Indices for Sugarcane

- **Normalised Difference Vegetation Index (NDVI)**

NDVI is normalized vegetation indices and it depicts the greenness of the vegetation and the values between 0.5 to 1 shows dense vegetation and less than that are sparse vegetation. It uses NIR and red bands of analysis. The NDVI indicator is extensively used to characterize vegetation covers due to its close relationship with several vegetation parameters such as Leaf Area Index (LAI).

In the early stages of planting, the soil still dominates the sugarcane. The values of the vegetation index NDVI are extremely low during this phase. During the second phase, the plant index values start to rise, but they are still low because they are still mixed with the soil spectral. As the vegetation index values continue to rise, they will peak in the third phase. The vegetation index values dropped in the fourth phase. In this state, the sugar cane plant is prepared for harvest. The values of the vegetation index drop significantly after harvest. This phenomenon is helpful for tracking sugarcane plantations because it happens repeatedly.

- **Enhanced Vegetation index (EVI)**

The vegetation index that is more sensitive in areas with thick canopy is the Enhanced Vegetation Index (EVI). It is an enhancement of the NDVI that lessens the impact of the atmosphere.

It uses NIR, Red and blue bands for analysis. The Enhanced Vegetation Index map differs from the NDVI despite having a similar appearance. The values in this case range from 0 to 0.6 (1 d.p.), indicating that areas that were previously regarded as extremely dense in the NDVI are now seen as less dense. Most helpful in areas with strong Leaf Area Index (LAI) where the NDVI may get saturated. Vegetation pixels should have EVI values between 0 and 1.

The anomalies in pixel values in an EVI image can be caused by both dark things, like water, and bright features, such clouds and white structures. It is employed to evaluate the variability of crop development under both sparse and thick vegetation cover conditions.

- **Ratio Vegetation Index (RVI)**

It is used for sparse vegetation. It uses NIR and red bands. NDVI gets saturated at grand growth stages and so the novel method can be EVI and RVI are used along with NDVI for dense and sparse categorisation. The Ratio vegetation index is 1 or nearly 1 if the reflectance's of the RED and NIR bands, are the same or comparable. For bare soils, S values are typically close to 1. The RVI raises as a pixel's amount of green vegetation grows. The RVI indicates if sparse vegetation is there and helps the farmers to diagnose the reason behind the sparse vegetation and is it caused by soil or slopes in the plot where the water is not retained which lead to sparse vegetation.

1. Calculate the NDVI using the formula

$$NDVI = \frac{(\rho NIR - \rho R)}{(\rho NIR + pR)} \tag{4}$$

i.e.: B08 – B04 / B08 + B04

ρNIR is the reflectivity of infrared band, ρR is the reflectivity of red band, ρB is the reflectivity of blue band.

B08 is the NIR band in Sentinel-2 imagery

B04 is the Red band in Sentinel-2 imagery

B02 is the Blue band in Sentinel-2 imagery

2. Calculate the EVI using the formula

$$2.5 \times \left(\frac{(\rho NIR - \rho R)}{(\rho NIR + 6 \times \rho R - 7.5 \times \rho B + 1)}\right) \tag{5}$$

i.e.: "2.5 × ((B08 – B04) / (B08 + 6 × B04 – 7.5 × B02 + 1))"

3. Calculate the RVI using the formula

$$\frac{\rho NIR}{\rho Red} \tag{6}$$

i.e.: "B08 / B04"

B08 is NIR band in Sentinel-2 imagery

B04 is Red band in Sentinel-2 imagery

B02 is Blue band in Sentinel-2 imagery

- **Need of NDVI, EVI and RVI for sugarcane crop**

Since sugarcane differs from many other crop types in terms of growth pattern and phenology, statistical and ML techniques may be used to analyse the spectral and temporal properties of satellite data in order to more effectively distinguish sugarcane.

The two types of agricultural practices done in sugarcane are seedling and other is ratoon sugarcane. The part of the plant cane which is kept after harvesting is done and again that crop is grown and it is known as ratoon cane .The ratoon sugarcane matures faster as compared to plant sugarcane and requires 11 months as compared to plant cane which requires 12 to 15 months. The analysis requires more study on identifying sugarcane varieties, determining the difference between plant and ratoon canes, and classifying sugarcane phenology. The health status of the crop and its checking periodically is necessary for crop production. The need for indices as RS is done and the satellite imagery spectral bands are explored along with the vegetation indices for the monitoring of the crop and categorizing into dense and sparse vegetation EVI helps in dense canopy as it removes the atmospheric effects and when the crop matures as the leaves of sugarcane grows. NIR reflectance varies and becomes less and NDVI results are not satisfactory so Enhanced vegetation index helps in predicting the harvest .The maximum NDVI gives the ready to grow sugarcane growth .The sparse vegetation value shows the soil is not good for growth of sugarcane or other factors like disease can be diagnosed for the sugarcane crop and so the farmer can take care of it as individual month is monitored for sugarcane crop. The quality of the crop can be monitored by dense and sparse vegetation values as shown in Tables III–V.

TABLE III. NDVI TIF (PYTHON) FOR BELGAVI REGION

Month	Year	NDVI Range	Average of Sparse Vegetation	Average of Dense Vegetation	Average of Full Vegetation
January	2022	−0.6586 to 0.8519	0.310604072	0.615623869	0.554064943
February	2022	−0.5391 to 0.8603	0.314759895	0.618865703	0.529428469
March	2022	−0.4405 to 0.9329	0.315056459	0.627221234	0.543019876
April	2022	−0.5420 to 1	0.315116996	0.616281179	0.532516541
May	2022	−0.4124 to 1	0.316403345	0.64832272	0.578262583
June	2022	−0.1814 to 0.8263	0.315931892	0.652012004	0.582477468
July	2022	−0.0465 to -	0.31419263	0.619816485	0.51294
August	2022	−0.5412 to 0.7675	0.316445425	0.642802597	0.570127184
September	2022	−1 to 1	0.316100818	0.683009268	0.616737074
October	2022	−1 to 0.9890	0.31613479	0.675496103	0.614734861
November	2022	−1 to 0.9485	0.316008693	0.643167178	0.571201521
December	2022	−1 to 0.9974	0.315741575	0.612179155	0.529843242
January	2023	−1 to 1	0.315009669	0.611499897	0.52875926
February	2023	−1 to 0.9106	0.314476405	0.617056245	0.52795217
March	2023	−1 to 0.9781	0.315772069	0.619630533	0.535878661

TABLE IV. RVI FOR BELGAVI REGION

Month	Year	Average of Sparse Vegetation	Average of Dense Vegetation	Average of Full Vegetation
June	2022	0.313446938	0.576569781	0.515440232
July	2022	0.312669907	0.579305644	0.52297797
August	2022	0.316322992	0.568555323	0.512235504
September	2022	0.312820547	0.585289827	0.52384459
October	2022	0.314652837	0.584413132	0.526253803
November	2022	0.319186107	0.571260985	0.520104613
December	2022	0	0.666666667	0.666666667
January	2023	0.314296683	0.597376173	0.528299359
February	2023	0.312046473	0.583261668	0.53506842
March	2023	0.326992609	0.586618725	0.544354008

TABLE V. EVI FOR BELGAVI REGION

Month	Year	Average of Sparse Vegetation	Average of Dense Vegetation	Average of Full Vegetation
January	2022	0.305205191	0.59565037	0.492381732
February	2022	0.305065227	0.583117181	0.464671912
March	2022	0.305024853	0.590855202	0.476094558
April	2022	0.305304259	0.620442483	0.52577327
May	2022	0.307420439	0.670312499	0.67025864
June	2022	0.306545723	0.702404908	0.702362748
July	2022	0.305973013	0.653092229	0.584002017
August	2022	0.306424127	0.755298957	0.755229918
September	2022	0.305567965	0.638286562	0.544051889
October	2022	0.305208223	0.641330805	0.548291481
November	2022	0.305114667	0.609410013	0.508291357
December	2022	0.305637031	0.649558903	0.586241406
January	2023	0.305106183	0.582670627	0.470853453
February	2023	0.304487428	0.581353651	0.460050408
March	2023	0.305031294	0.58623834	0.471225664

IV. RESULT AND DISCUSSION

The experiments are conducted using the Jupyter notebook standards as the basis, using Google Earth Engine and Google Colaboratory platforms. Google Colaboratory facilitates the continuous use of GPU and GPU hardware accelerators, and permits code storage in Google Drive. Geospatial big data technologies such as Google Earth Engine are used by Python 3 and Google Colaboratory. For the pre-processing done using Python the steps are:

A. Preprocessing on Satellite Data

1. The initial preprocessing is done on the data given by Belagavi.
2. The geographic coordinates (latitude and longitude) were taken from the farmers as ground truth and convert the farmer's jpg image to a kml file using a python code to get the farm placeholder.
3. Put the kml file path in the merged coordinate's python code to get the coordinates of the whole area.
4. The validation was done for the merged area.
5. To get the datasets from the Sentinel 2 dataset based on the merged coordinates.
6. Sentinel-2 satellite imagery is fetched for a specific location and time range, selects specific bands of interest, and exports the processed image to Google Drive as a GeoTIFF file for further analysis as shown in the shape file of Fig. 3.
7. Select the image bands.
8. Define the visualization parameters for display.
9. Apply overlays to fuse / merge multiple images together.
10. These fused images can further be used to apply ML models.

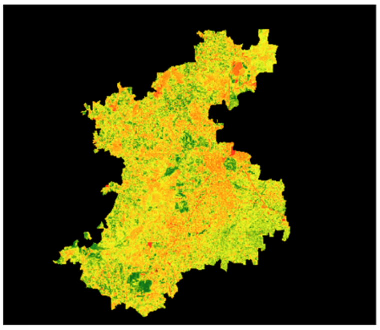

Fig. 3. Belagavi shape file.

B. Observations of the Tabular data NDVI, EVI and RVI from Tables III–V

1) Impact of dense and sparse vegetation from Tables III–V

1. The cutting period matches the ground truth as per the tabular results in Table III and December month is the cutting period observed from average NDVI values.

2. As the sugarcane crop starts maturity as shown from NDVI values in September and then NDVI drops due to canopy growth and when there is dense canopy EVI gives better results as NDVI values get saturated and NIR

bands reflectance is low and caused low NDVI values. EVI values can be used then when the sugarcane is matured for analysis.

3. Maximum NDVI for Belagavi can be seen around 1. RVI is used for sparse vegetation.

4. The ratoon sugarcane growth is in average for 11 months and fast as shown from the tabular data. The second ratoon sugarcane growth is fast as compared to plant cane and first ratoon sugarcane. In Belagavi region the soil is good for the sugarcane crop to grow .The plant cane grows in 8 months as shown from Tables III–V and the first ratoon sugarcane grows in 6 months and further second ratoon will be faster as compared to first ratoon.

5. The average dense vegetation shows the plot area is good in terms of soil, water for the sugarcane to grow.

6. The sparse vegetation indicates area has slopes and water is not retained and creates problems in the growth of the sugarcane.

7. NDVI, EVI and RVI graphs are plotted for the years 2021 to 2023.

8. June, July there is cloud cover and NDVI is affected a seen from tabular data.

9. The date of plantation and emergence period can be found by monitoring the growth.

10. The vegetation density is calculated for some plots of the farmers for the year 2022 to 2023 of the same region as shown in Table VI which helps the farmers to understand the dense and sparse vegetation in their farms and accordingly work for the development of the sugarcane growth and diagnose the causes of sparse vegetation.

The slope of land or water retention issues was detected through the vegetation indices. NDVI results were low when it reached maturity and for dense canopy the results of EVI were seen to be more .The text marked in red for Tables III–V is the maturity period of the sugarcane crop.

The vegetation density as shown in Table VI for plot 3 is seen to be more of sparse vegetation and when discussed with the farmers it was validated that there were slopes in the land and water was not retained which lead to sparse vegetation.

TABLE VI. VEGETATION DENSITY OF THE PLOTS

Plot No.	Year	Sparse percentage	Dense Percentage	Area of the plot
1	2021	18.2934	37.3786	3 acres
	2022	18.764	37.8673	
	2023	18.7667	37.9413	
2	2022	18.7882	37.847	3 acres
	2023	18.8499	37.8879	
3	2022	27.9931	56.7156	2 acres
	2023	28.2502	56.6307	
4	2023	20.4204	40.1512	3 acres
5	2023	38.1217	75.9676	1.5acres
6	2021	5.89747	20.013	5 acres
	2022	6.03014	20.0321	
	2023	6.00518	20.0601	
7	2022	19.9034	35.531	3 acres
	2023	18.824	36.2925	
8	2022	22.3994	22.3994	5 acres
	2023	22.7439	22.7439	

The plot 1 was good as the percentage of dense vegetation is growing in the years from 2021 to 2023 as shown from the Table VI and the soil is good for sugarcane growth. The indices categories dense and sparse vegetation which will help the farmers to diagnose the reason behind sparse vegetation. So monitoring monthly results in good agriculture practices. Moreover ML models when deployed on the vegetation indices further help in getting the accuracy ,precision and recall for the plots collectively. Also barren, water and land for the plots was categorized using the ML models like SVM, RF, Maximum likelihood, KNN and ISO clustering as shown in Tables VII–XI.

TABLE VII. EVALUATION METRICS FOR SVM CLASSIFIER

Plot Categorisation	Precision	Recall	F1-Score
Barren land	0	0	0
Built up	1	0.09	0.24
Dense	0.96	0.90	0.94
Shrub and grassland	0.79	0.84	0.81
Sparse	0.84	0.99	0.91
Water	1	0.86	0.91
Accuracy		**0.87**	
Macro avg	0.85	0.73	0.72
Weighted avg	0.82	0.84	0.81

If more hyper spectral dataset or multispectral dataset is publicly available in future then not only monitoring but all aspects of agriculture in terms of yield, crop management and disease detection will be resolved. The insights from the study can be applied in real world agricultural practices and will benefit the farmers in terms of pesticides for their own plots as the study shows dense and sparse vegetation. The sparse vegetation causes can be then diagnosed by the farmers well in time before maturity of the sugarcane crop leading to more sugar and ethanol. The sugarcane mills will be benefited also as the production of sugarcane will grow more.

C. ML Analysis

SVM classification results are better than random forest, maximum likelihood, K Nearest Neighbor, and ISO clustering. SVM classification is a supervised learning model used for regression and classification. Table III, shows the visualization of ML models applied on the image with cloud cover removed. The RGB color space image and the output give a proper visualization of water, barren land, and cultivated land after applying the SVM classifier. The colors depict the various classes of SVM. Under the premise that the statistics of each class in each band are regularly distributed, maximum likelihood classification can provide an estimate of the probability that a particular pixel belongs to a specific class. Pixels are then placed in the most likely category. It shows that the percentage of land that is classified as barren is higher than any of the Random Trees classification method is a supervised ML classifier that builds numerous decision trees, assigns each tree a random subset of factors, and uses the most common tree output to determine the classification as a whole. The random tree classification gives better results than the maximum likelihood and will perform on the random selection of training pixels. The vegetation class is not visualized totally as compared to the SVM classifier. The water and barren land are classified to some extent as per the original raster image. K nearest neighbor stores all the given input data and at the time of classification it performs the actions on the dataset. The K value can be 1, 2, 3 or any and depending on K nearest neighbors the classification is done. The classification of the 3 classes is done. The dataset if it is more than KNN would be more effective.

Download the image TIF and add an image in Google Earth engine after preprocessing and making cloud cover 0. The zoomed input image of Copernicus is better as compared to Earth Explorer USGS satellite imagery as visualization can be done effectively and due to the Near Infrared band and Red band index, the vegetation index can be found using the equation (NIR-Red) / (NIR + Red) and sugarcane plots can be accessed. Below Fig. 4, NDVI of negative value approaching -1 shows water and a value of 0.5 indicates vegetation. The output indicates scanty vegetation, which agrees the assumption that healthy thick vegetation has a value greater than 0.5.

The work is done on Sentinel-2 imagery from Earth Explorer and Sentinel-2 Multispectral Copernicus imagery. The results with cloud cover removed gave good interpretation about the classes like vegetation, barren, and water after classifying with ML techniques. In the image with cloud cover from Earth Explorer, the results are not interpreted properly; therefore, Copernicus satellite imagery is taken for better results and interpretation .The ML techniques were applied to remote sensed data and SVM classification as shown in evaluation Table VII gave better results as compared to random trees, maximum likelihood, ISO clustering, and K nearest Neighbor classifiers. The zoomed input image of Copernicus is better than Earth Explorer USGS satellite imagery as visualization can be done effectively. The Near Infrared band and Red band index can be found for Copernicus data and classes are separated effectively using NDVI and ML techniques. The evaluation metrics used in this paper are as follows:

Precision: It is the proportion of all expected positives to accurately forecasted positive observations.

Recall: It's the proportion of accurately anticipated positive observations to real positive ones.

F1-Score: It is the precision and recall harmonic mean.

Accuracy: It is the proportion of all observations to accurately predicted observations.

Macro Average: It determines the average performance throughout all classes, weighing each one equally.

Weighted Average: It determines the average performance for each class, although it gives classes with more instances greater weight. When classes are unbalanced, it is helpful.

The evaluation metrics shown for SVM classifier in Table VII as the visualization after applying ML models to the image with cloud cover removed is better and the accuracy is 87% and other evaluation metrics are precision, recall, F1-Score and Support. As shown in Table VIII the accuracy is 81% as Random trees does the random assignment of pixels and so the visualization is not good as compared to SVM classifier. As shown in Table IX, the KNN accuracy is 82% and is better than Random trees classification and based on value of K.

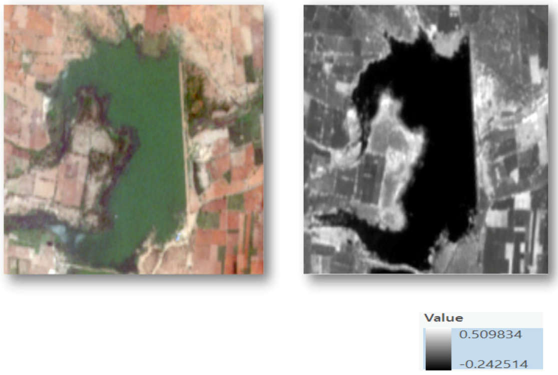

Fig. 4. NDVI analysis.

TABLE VIII. EVALUATION METRICS FOR RANDOM TREE

Plot Categorisation	Precision	Recall	F1-Score
Built up	1	1	1
Dense	1	0	0
Sparse	0.70	1	0.85
Accuracy		0.81	
Macro avg	0.91	0.66	0.61
Weighted avg	0.84	0.81	0.70

TABLE IX. EVALUATION METRICS FOR KNN

Plot Categorisation	Precision	Recall	F1-Score
Built up	0.40	1	0.68
Dense	0.91	1	0.89
Shrub and grassland	0.74	1	0.85
Sparse	1	0.67	0.74
Water	1	1	1
Accuracy		0.82	
Macro avg	0.86	0.90	0.85
Weighted avg	0.88	0.82	0.82

As shown in Table X, ISO clustering accuracy is not good as it is unsupervised algorithm and the classes are not visualised precisely. So in further work more focus should be made on unsupervised techniques also.

As shown in Table XI, the maximum likelihood evaluation parameters on the image with cloud cover removed is shown. The ML models were applied for separation of the classes like water, barren, built up area and vegetation. ML models can be applied to the vegetation indices for dense and sparse vegetation in the individual plot of the farmer for diagnosing the vegetation each month. The ML models can also estimate the quality indicators of satellite imagery data in further work. The NDVI values can be enhanced after atmospheric correction, especially for cloud-covered images, and further research is needed, especially during rainy seasons. Also, the enhanced vegetation indices along with NDVI can explore grand growth stages of sugarcane where NDVI gets saturated, and ML models can be applied on the enhanced vegetation indices to get accurate information of dense canopy. Moreover, deep learning models will enhance the categorization more with hyper spectral data or large multispectral data of sugarcane crop. There is more work to be done in the field of agriculture, as the use of technology for self-supervised learning on small datasets has not got much attention. Though data augmentation, suitable assessment criteria, and numerous model ensembles are essential for small samples and should be investigated in the RS domain in agriculture, few-shot and zero-shot learning techniques can be efficient for optimizing performance from limited data.

TABLE X. EVALUATION METRICS FOR ISO CLUSTERING

Plot Categorisation	Precision	Recall	F1-Score
Built up	0.5	0.5	0.5
Sparse	0.5	0.5	0.5
Water	0	1	0
Accuracy		0.5	
Macro avg	0.26	0.5	0.34
Weighted avg	0.33	0.5	0.39

TABLE XI. EVALUATION METRICS FOR MAXIMUM LIKELIHOOD

Plot Categorisation	Precision	Recall	F1-Score
Built up	0.73	0.71	0
Sparse	1	1	1
Water	0	1	0
Accuracy		0.72	
Macro avg	0.71	0.70	0.70
Weighted avg	0.89	0.67	0.67

As shown in Table XII the SVM results in terms of accuracy are better than other models and helps in categorization of vegetation effectively. As the dataset is limited and so if the large dataset for satellite imagery is found then the efficiency to categories the RS data will definitely increase in terms of accuracy, precision and Recall

TABLE XII. EVALUATION TABLE OF ML CLASSIFIERS

Models	Accuracy	Precision	Recall
SVM	87	85	73
Random Trees	81	91	66
KNN	82	86	90
Maximum likelihood	72	71	70
ISO clustering	50	25	50

D. Challenges

From the literature survey done, the Leaf Area Index (LAI) is an effective method for tracking crop development; although the current methodology needs to be expanded for large areas and more parameters need to be researched for LAI. In the case of low and high crop coverage, the vegetation indices need to be explored.

For all weather conditions, RS and GIS should be used for better results. Temperature, precipitation, and soil moisture can be added to generalize the model.

Multidate satellite data sets with higher spatial resolutions are needed to study sugarcane crop growth and crop stresses. Understanding the crops spectral features, it requires gathering ground truth data for correlation with satellite imagery. The exploration of suitable vegetation indices for crop growth monitoring needs to be worked more for better analysis of sugarcane crops as research shows that NDVI is highly susceptible to soil backdrops. To increase precision in agricultural yield estimation, it would help calculate LAI correlation with other vegetation indices across different stages of crop development. Multi-crop cultivation work and its phenology need to be explored more.ML models can be explored for the same.

After atmospheric and geometric corrections, the relationship between crop growth stages and sugarcane yield needs further study. Crop stress due to disease or water stress or nutrient deficiency detection is required for crop assessment. ML and deep learning techniques should be explored more for the classification of the crop. The analysis requires more study on identifying sugarcane varieties, determining the difference between plant and ratoon canes, and classifying sugarcane phenology. The accuracy of crop yielding needs to be improved and the analysis of intercrop yield needs to be explored more.

After atmospheric correction, the NDVI values can be improved particularly for images with cloud cover. Further work needs to be explored more on satellite images with cloud cover, especially during the rainy season. The vegetation indices using satellite imagery can help identify dense vegetation for sugarcane growth, while sparse vegetation indicates slopes and water issues. The deep learning and ML techniques can further be used on the vegetation indices to find the best harvest time for the farmer.

E. Future Prospects

The analysis of remotely sensed data of sugarcane crops using ML methods should be focused more as ensemble ML methods can serve the purpose. The satellite imagery data can provide better results on dense and sparse canopy for farmers, allowing for better prediction of harvest dates and identifying gaps in plantation and emergence periods. The enhanced vegetation indices can diagnose second-ratoon sugarcane growth. However, challenges include identifying dense vegetation for individual farmers and sparse vegetation for slopes and water retention issues. Implementing ML algorithms requires significant processing power, although the application of RS can be enhanced by algorithmic developments that shorten calculation times.

The hyper spectral imagery data if available worldwide will enhance the use of ML models and the outcome will be more promising. The cutting period of the sugarcane should match with the indicators of satellite data and for that ground truth data should be also collected. In the June and July months, there is cloud cover so atmospheric correction done by ML methods will give better NDVI and help in dense canopy calculation. SVM classifier along with RF classifier if used in Deep learning models as a feature extractor will give better results in terms of accuracy. From the survey ML models have not been used with full potential on the enhanced vegetation indices to get dense canopy. Further satellite imagery data along with meteorological data will aid in more results on dense or sparse canopy of sugarcane and its effects on the crop can be monitored well. The spectral bands should be explored more in-depth as spatial information along with temporal and spectral information will help in monitoring the intercrop yield also. The dataset should be of high resolution and for that hyper spectral imagery data of sugarcane crops in India should be available for further analysis. The relocation matching also should be precise for monitoring each farm individually. ML techniques should be used more for ratoon cane as ratoon cane grows faster as compared to plant cane. The unsupervised classification techniques on satellite imagery data should be also explored.

V. Conclusions

RS and GIS technologies were studied in the agriculture domain. RS of various platforms was studied in the literature survey and work was done on Sentinel-2 imagery from Earth Explorer and Sentinel-2 Multispectral Copernicus imagery and the results with cloud cover removed gave good interpretation about the classes like vegetation, barren, and water after classifying with ML techniques. In the image with cloud cover from Earth Explorer, the results are not interpreted properly; therefore, Copernicus satellite imagery is taken for better results and interpretation is better. The ML techniques were applied to remotely sensed data, and SVM classification gave better results as compared to random trees, maximum likelihood, ISO clustering, and K nearest Neighbor classifiers. The zoomed input image of Copernicus is better than Earth Explorer USGS satellite imagery as visualization can be done effectively and due to the Near Infrared band and Red band index the vegetation index can be found and classes are separated effectively using RS and ML techniques. The satellite imagery data, with vegetation indices, can provide better remote results on dense and sparse canopies if monitored monthly, reducing the impact of dust and atmospheric effects on farms. The NDVI values can be enhanced after atmospheric correction, especially for cloud-covered images, and further research is needed, especially during rainy seasons. The novel method is to use an indicator that is less susceptible to leaf and canopy structure than NDVI but is substantially associated with leaf chlorophyll concentration during the extensive growth stage. In addition to normalized difference vegetation index (NDVI), other indices, enhanced vegetation index like EVI and RVI, if meteorological data, is used in the future for sugarcane analysis and proper monitoring. The ratio vegetation indices serve sparse needs, while enhanced vegetation indices provide better results in dense canopy. The ML models are applied to vegetation indices to find the dense canopy and sparse canopy of the sugarcane crop in the same plot remotely which will help the farmer as the farmer can take the steps for the sparse vegetation before time. The ML models are applied on the tabular data of NDVI +EVI+RVI and helps in categorizing the barren, land and water and also evaluating based on Precision, Recall and Accuracy. The combination of vegetation indices and ML models can enhance the categorization of the sugarcane crop effectively for dense and sparse vegetation leading to precision agriculture and sustainably too. In further work ML models can be also applied on the images of NDVI, EVI and RVI. The fusion of tabular and image data can be used for further analysis of dense and sparse vegetation.

The requirement for significant processing power and the high cost of ML systems remain obstacles, too. The research can be focused on causes of sparse vegetation and dense vegetation using Deep learning methods. The deep learning techniques can effectively manage water stress levels in sugarcane crops, particularly in areas with sparse vegetation. The further analysis of NDVI for multi-crop assessment can be done also.

Remotely sensed sugarcane crop data can be analyzed using ML techniques to reveal information on canopy, harvest dates, and plantation date gaps. The difficulties, however, include determining the best vegetation for individual farmers as well as resolving the problems of water retention and slope vegetation. Hyper spectral data can improve the usage of deep learning models, and algorithmic advancements can improve the use of RS. Ratoon sugarcane, which is grown for a year after the plant cane is harvested, can be further investigated using satellite imagery data, and the sugarcane cutting period can correspond with satellite data indications. RS and deep learning can be used to achieve the same in future.

To investigate sugarcane crop growth and stress, multidate satellite data sets with increasing geographical resolutions are required. More research is required to define sugarcane phenology, identify sugarcane types, and distinguish between plant and ratoon cane. The atmospheric correction and cloud-free satellite imagery are also required for analysis. Also ground truth data from the farmers will help in mapping the georeferenced satellite data.

Additional spectral bands, meteorological information, and satellite imagery data can be used to track intercrop yield and enhance crop management. The results can be improved and deep learning techniques can be applied more effectively with atmospheric adjustment for satellite photography data. Sensitivity analysis can be used as a feature selection method, selecting the least significant feature to eliminate from each backward selection cycle, and visualisation tools can be used to further investigate the spectral characteristics.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

The original draft preparation, design, analysis, and investigation were done by Mansi Kambli (corresponding author). The work was reviewed by Dr. Bhakti Palkar. The manuscript was read and approved by all authors.

ACKNOWLEDGMENT

I express my deep gratitude to K. J. Somaiya College of Engineering, Somaiya Vidyavihar University for their support. I also thank the Copernicus and Earth Explorer sites for the dataset.

REFERENCES

[1] S. K. Singla, R. D. Garg, and O. P. Dubey, "ML models to estimate the sugarcane brix values from multitemporal vegetation indices," in *Proc. 2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM)*, IEEE, 2020, pp. 177–183. https://doi.org/10.1109/ICCAKM46823.2020.9051545

[2] J. Wellens, M. Stasolla, M. T. Sall, B. Tychon, and X. Neyt, "Assimilation of Sentinel-1 change detection in the aquacrop model: Case of sugarcane," in *Proc. 2021 IEEE Int. Geosci. Remote Sens. Symp., IGARSS*, 2021, pp. 1069–1072. https://doi.org/10.1109/IGARSS47720.2021.9555163

[3] X. Yuan *et al.*, "Monitoring of sugarcane crop based on time series of Sentinel-1 data: A case study of Fusui, Guangxi," in *Proc. 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, IEEE, 2019. https://doi.org/10.1109/Agro-Geoinformatics.2019.8820221

[4] H. Fang and Y. Wang, "Crop yield estimation at field scales by assimilating time series of Sentinel-2 Data into a modified CASA-WOFOST coupled model," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022. https://doi.org/10.1109/TGRS.2020.3047102

[5] H. Lin, J. Chen, Z. Pei, S. Zhang, and X. Hu, "Monitoring sugarcane growth using ENVISAT ASAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2572–2580, 2009. https://doi.org/10.1109/TGRS.2009.2015769

[6] L. Martinez-Ferrer, M. Piles, and G. Camps-Valls, "Crop yield estimation and interpretability with gaussian processes," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 12, pp. 2043–2047, 2021. https://doi.org/10.1109/LGRS.2020.3016140

[7] E. J. Jones, T. F. A. Bishop, B. P. Malone, P. J. Hulme, B. M. Whelan, and P. Filippi, "Identifying causes of crop yield variability with interpretive ML," *Computers and Electronics in Agriculture*, vol. 192, 106632, 2022. https://doi.org/10.1016/j.compag.2021.106632

[8] S. S. Virnodkar, V. K. Pachghare, V. C. Patil, and S. K. Jha, "DenseResUNet: An architecture to assess water-stressed sugarcane crops from Sentinel-2 satellite imagery," *Traitement du Signal*, vol. 38, no. 4, 2021. https://doi.org/10.18280/ts.380424

[9] S. S. Virnodkar, V. K. Pachghare, V. C. Patil, and S. K. Jha, "Application of ML on RS data for sugarcane crop classification: A review," *ICT Analysis and Applications*, pp. 539–555, 2020.

[10] S. S. Virnodkar, V. K. Pachghare, V. C. Patil, and S. K. Jha, "CaneSat dataset to leverage convolutional neural networks for sugarcane classification from Sentinel-2," *Journal of King Saud University-Computer and Information Sciences*, 2020. https://doi.org/10.1016/j.jksuci.2020.09.005

[11] S. M. Acharya, S. S. Pawar, and N. B. Wable, "Application of RS GIS in agriculture," *International Journal of Advanced Engineering Research and Science*, vol. 5, no. 4, 237434, 2018. https://dx.doi.org/10.22161/ijaers.5.4.10

[12] J. Useya and S. Chen, "Comparative performance evaluation of pixel-level and decision-level data fusion of Landsat 8 OLI, Landsat 7 ETM+ and Sentinel-2 MSI for crop ensemble classification," *IEEE Journal of Selected Topics in Applied Earth Observations and RS*, vol. 11, no. 11, pp. 4441–4451, 2018. https://doi.org/10.1109/JSTARS.2018.2870650

[13] Z. Li and Z. Chen, "RS indicators for crop growth monitoring at different scales," in *Proc. 2011 IEEE International Geoscience and RS Symposium*, IEEE, 2011, pp. 4062–4065. https://doi.org/10.1109/IGARSS.2011.6050124

[14] S. Khanal *et al.*, "Remote sensing in agriculture—Accomplishments, limitations, and opportunities," *Remote Sensing*, vol. 12, no. 22, 3783, 2020.

[15] S. Virnodkar, V. K. Pachghare, V. C. Patil, and S. K. Jha, "Performance evaluation of RF and SVM for sugarcane classification using Sentinel-2 NDVI time-series," *Progress in Advanced Computing and Intelligent Engineering*, pp. 163–174. Springer, Singapore, 2021.

[16] S. Supavetch, "In-season yield prediction and monitoring of sugarcane using cumulative growth of normalized difference vegetation index," in *Proc. 2022 IEEE Mediterranean and Middle-East Geoscience and RS Symposium (M2GARSS)*, IEEE, 2022, pp. 173–176. https://doi.org/10.1109/M2GARSS52314.2022.9840038

[17] M. P. Mello, C. Atzberger, and A. R. Formaggio, "Near real-time yield estimation for sugarcane in Brazil combining RS and official statistical data," in *Proc. 2014 IEEE Geoscience and RS Symposium*, IEEE, 2014, pp. 5064–5067. xhttps://doi.org/10.1109/IGARSS.2014.6947635

[18] S. S. Virnodkar, V. K. Pachghare, V. C. Patil, and S. K. Jha, "RS and ML for crop water stress determination in various crops: A critical review," *Precision Agriculture*, vol. 21, no. 5, pp. 1121–1155, 2020. https://doi.org/10.1007/s11119-020-09711-9

[19] R. N. Sahoo, S. S. Ray, and K. R. Manjunath. (2015). Hyperspectral remote sensing of agriculture. *Current Science*. [Online]. 108(5), pp. 848–859. Available: http://www.jstor.org/stable/24216514

[20] C. Ashcraft and K. Karra, "ML aided crop yield optimization," arXiv preprint, arXiv: 2111.00963, 2021. https://doi.org/10.48550/arXiv.2111.00963

[21] S. Tang, Q. Zhu, X. Zhou, S. Liu, and M. Wu, "A conception of digital agriculture," *IEEE International Geoscience and RS Symposium*, vol. 5, pp. 3026–3028, 2002. doi: 10.1109/IGARSS.2002.1026858

[22] P. S. M. Gopal, "Performance evaluation of best feature subsets for crop yield prediction using ML algorithms," *Applied Artificial Intelligence*, vol. 33, no. 7, pp. 621–642, 2019. https://doi.org/10.1080/08839514.2019.1592343

[23] S. Deng, M. Gao, C. Ren, S. Li, and Y. Liang, "Extraction of sugarcane planting area based on similarity of NDVI time series," *IEEE Access*, vol. 10, pp. 117362–117373, 2022. https://doi.org/10.1109/ACCESS.2022.3219841

[24] R. R. do Valle Gonçalves, J. Zullo, T. M. Peron, S. R. M. Evangelista, and L. A. S. Romani, "Numerical models to forecast the sugarcane production in regional scale based on time series of NDVI/AVHRR images," in *Proc. 2015 8th International Workshop on the Analysis of Multitemporal RS Images (Multi-Temp)*, IEEE, 2015, pp. 1–4. https://doi.org/10.1109/Multi-Temp.2015.7245806

[25] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International Journal of RS*, vol. 28, no. 5, pp. 823–870, 2007. https://doi.org/10.1080/01431160600746456

[26] J. Som-ard, C. Atzberger, E. Izquierdo-Verdiguier, F. Vuolo, and M. Immitzer, "RS applications in sugarcane cultivation: A review," *Remote Sens.*, vol. 13, no. 20, 4040, 2020. https://doi.org/10.3390/rs13204040

[27] E. Panwar, A. N. J. Kukunuri, D. Singh, A. K. Sharma, and H. Kumar, "An efficient ml enabled non-destructive technique for remote: Monitoring of sugarcane crop health," *IEEE Access*, vol. 10, pp. 75956–75970, 2022. https://doi.org/10.1109/ACCESS.2022.3190716

[28] P. M. Kai, R. M. da Costa, B. M. de Oliveira, D. S. A. Fernandes, J. Felix, and F. Soares, "Discrimination of sugarcane varieties by RS: A review of the literature," in *Proc. 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, IEEE, 2020, pp. 1212–1217.

[29] M. R. Jamnani, A. Liaghat, and F. Mirzaei, "Optimisation of sugarcane harvest USI remote sensing," *International Archives of the Photogrammetry, RS & Spatial Information Sciences*, 2019.

[30] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018. https://doi.org/10.1016/j.compag.2018.02.016

[31] B. Victor, Z. He, and A. Nibali, "A systematic review of the use of deep learning in satellite imagery for agriculture," arXiv preprint arXiv: 2210.01272, 2022. https://doi.org/10.1109/JSTARS.2024.3501216

[32] M. K. Villareal and A. F. Tongco, "RS techniques for classification and mapping of sugarcane growth," *Engineering,*

Technology & Applied Science Research, vol. 10, no. 4, pp. 6041–6046, 2020.

[33] M. Yeasin, D. Haldar, S. Kumar, R. K. Paul, and S. Ghosh, "ML techniques for phenology assessment of sugarcane using conjunctive SAR and optical data," *Remote Sens.*, vol. 14, no. 14, 3249, 2022. https://doi.org/10.3390/rs14143249

[34] T. F. Canata, M. C. F. Wei, L. F. Maldaner, and J. P. Molin, "Sugarcane yield mapping using high-resolution imagery data and ML technique," *Remote Sens.*, vol. 13, 232, 2021. https://doi.org/10.3390/rs13020232

[35] R. Scrivani, J. Zullo, and L. A. S. Romani, "SITS for estimating sugarcane production," in *Proc. 2017 9th International Workshop on the Analysis of Multitemporal RS Images (MultiTemp)*, IEEE, 2017, pp. 1–4. https://doi.org/10.1109/Multi-Temp.2017.8035254

[36] E. M. Abdel-Rahman and Fethi B. Ahmed, "The application of RS techniques to sugarcane (Saccharum spp. hybrid) production: A review of the literature," *International Journal of RS*, vol. 29, no. 13, pp. 3753–3767, 2008. https://doi.org/10.1080/01431160701874603

[37] V. Kumar and Y. S. Rao, "Temporal analysis of different crops using quad-pol RADARSAT-2 data," in *Proc. 2015 IEEE International Geoscience and RS Symposium (IGARSS)*, IEEE, 2015, pp. 3211–3214.

[38] R. Ly, K. Dia, and M. Diallo, "RS and ML for food crop production data in Africa post-COVID-19," arXiv preprint, arXiv:2108.10054, 2021. https://doi.org/10.48550/arXiv.2108.10054

[39] K. Krupavathi, M. Raghubabu, A. Mani, P. R. K. Parasad, and L. Edukondalu, "Field-scale estimation and comparison of the sugarcane yield from RS data: A ML approach," *Journal of the Indian Society of RS*, vol. 50, no. 2, pp. 299–312, 2022. https://doi.org/10.1007/s12524-021-01448-w

[40] A. Mitra, S. L. T. Vangipuram, A. K. Bapatla, V. KVV Bathalapalli, S. P. Mohanty, E. Kougianos, and C. Ray, "Everything you wanted to know about smart agriculture," arXiv preprint, arXiv:2201.04754, 2022. https://doi.org/10.48550/arXiv.2201.04754

[41] Q. Xie, J. Dash, W. Huang, D. Peng, Q. Qin, H. Mortimer, and R. Casa *et al.*, "Vegetation indices combining the red and red-edge spectral information for leaf area index retrieval," *IEEE Journal of Selected Topics in Applied Earth Observations and RS*, vol. 11, no. 5, pp. 1482–1493, 2018. https://doi.org/10.1109/JSTARS.2018.2813281

[42] A. Sharifi, "Using sentinel-2 data to predict nitrogen uptake in maize crop," *IEEE Journal of Selected Topics in Applied Earth Observations and RS*, vol. 13, pp. 2656–2662, 2020. https://doi.org/10.1109/JSTARS.2020.2998638

[43] C. Lozano-Garzon, G. Bravo-Córdoba, H. Castro, G. González-Rodríguez, D. Niño, H. Nuñez, and C. Pardo *et al.*, "RS and ML modeling to support the identification of sugarcane crops," *IEEE Access*, vol. 10, pp. 17542–17555, 2022. https://doi.org/10.1109/ACCESS.2022.3148691

[44] A. Ammar, A. Koubaa, and B. Benjdira, "Deep-learning-based automated palm tree counting and geolocation in large farms from aerial geotagged images," *Agronomy*, vol. 11, no. 8, 1458, 2021. https://doi.org/10.3390/agronomy11081458

[45] D. Singh, S. Srivastava, and G. D. R. Guru, "On new trends in agriculture, environmental & biological sciences for inclusive development effect of climate change on sugarcane crop: A review," *Journal of Pharmacognosy and Phytochemistry*, vol. 6, pp. 255–261, 2020.

[46] Z. Mamatkulov, E. Safarov, R. Oymatov, I. Abdurahmanov, and M. Rajapbaev, "Application of GIS and RS in real-time crop monitoring and yield forecasting: A case study of cotton fields in low and high productive farmlands," in *Proc. E3S Web of Conferences*, vol. 227, 03001. EDP Sciences, 2021. https://doi.org/10.1051/e3sconf/202122703001

[47] K. R. Manjula, S. Jyothi, and S. A. K. Varma, "Digitizing the forest resource map using ArcGIS," *International Journal of Computer Science Issues (IJCSI)*, vol. 7, no. 6, 300, 2010.

[48] M. S. Simoes, A. R. Formaggio, J. C. N. Epiphanio, and C. C. Freitas, "Evaluation of Radarsat multitemporal data for crop discrimination in highly cultivated areas," in *Proc. IEEE 1999 International Geoscience and RS Symposium, IGARSS'99*, IEEE, 1999, vol. 2, pp. 840–842. https://doi.org/10.1109/IGARSS.1999.774459

[49] G. M. Gandhi, B. S. Parthiban, N. Thummalu, and A. Christy, "NDVI: Vegetation change detection using RS and GIS—A case study of Vellore District," *Procedia Computer Science*, vol. 57, pp. 1199–1210, 2015. https://doi.org/10.1016/j.procs.2015.07.415

[50] S. K. Singla, R. D. Garg, and O. P. Dubey, "Ensemble ML methods to estimate the sugarcane yield based on RS information," *Rev.d'Intelligence Artif.*, vol. 34, no. 6, pp. 731–743, 2020. https://doi.org/10.18280/ria.340607

[51] A. B. Ballo, A. Coulibaly, and M. Diaby, "Optimization of Mango leaf disease detection using a hybrid Mlapproach," *International Journal Computer Theory and Engineering*, vol, 16, no 3, pp. 94–103, 2024. http://dx.doi.org/10.7763/IJCTE.2024.V16.1358

[52] S. Bunyang *et al.*, "Self-supervised learning advanced plant disease image classification with SimCLR," *Advances in Computational Intelligence*, vol. 3, no. 5, 18, 2023. https://doi.org/10.1007/s43674-023-00065-z

[53] U.S. Geological Survey. (April 1, 2022). USGS Sentinel—2A [Online]. Available: https://earthexplorer.usgs.gov/

[54] Copernicus OpenAccessHub. (April 21, 2022) Sentinel2A. [Online]. Available: https://scihub.copernicus.eu/dhus/#/home