Integrating NDVI, Zonal Statistics, and Random Forest for Precision Agricultural Monitoring in the Mahalaxmi Kheda Region

Shivani S. Bhosle[1], Dr. Prapti D. Deshmukh[2]

[1]Research Fellow, Dr. G. Y. Pathrikar College of Computer Science and Information Technology, MGM University, Chhatrapti Sambhajinagr, Maharashtra, Email: shivanibhosle21@gmail.com, ORCID: 0009-0003-4050-7420
[2]Pricipal, Dr. G. Y. Pathrikar College of Computer Science and Information Technology, MGM University, Chhatrapti Sambhajinagr, Maharashtra, Email: prapti.research1@gmail.com, ORCID: 000-0002-3065-6057

Abstract
The aim of this study was to monitor and map various crop types and their locations in the Mahalaxmi Kheda region using geospatial tools. Harmonized Landsat and Sentinel (HLS) datasets provided multitemporal imagery, which underwent preprocessing steps like atmospheric correction and layer stacking. The Normalized Difference Vegetation Index (NDVI) was calculated to evaluate plant health, growth patterns, and stress conditions. The Random Forest algorithm was employed to classify primary land cover and crop types in the study area. An accuracy assessment validated the classification results. Zonal statistics were used to analyze class distribution and spectral responses across different zones. The findings indicated that NDVI effectively distinguished between healthy and stressed plants, while zonal statistics quantified these differences at the regional level. The Random Forest classifier successfully identified crop types such as cotton, sugarcane, and sweet lime, achieving satisfactory accuracy levels. The overall classification accuracy was 81.17%, with a kappa coefficient of 0.7539. This study highlights the effectiveness of integrating NDVI analysis, regional statistics, and Random Forest methods for crop identification and agricultural monitoring, offering a reliable and cost-effective approach for precision agriculture.

Keywords: Harmonized Landsat Sentinel (HLS), Random Forest ,NDVI, Crop classification, Zonal statistics, Agricultural monitoring, Vegetation indices

1. Introduction

Agriculture not only provides food for people, but is also a very effective way to eliminate extreme poverty and promote economic growth. Agriculture plays a crucial role in the Indian economy, with a cultivation history spanning nearly 10,000 years. Currently, India ranks second among the largest producers of agricultural products in the world. Approximately 60% of the population is engaged in agriculture or related occupations[1]. Timely, relevant, and reliable information is essential for effective planning of the agricultural sector. Given the significant role of crop insurance in agricultural planning, regular monitoring of agricultural land is imperative. It is crucial to evaluate crop yields, acreages, and types at the block or tehsil level[2].The Indian Commission for Agriculture recognizes the important role of remote sensing (RS) technology in generating high-quality agricultural data. India was among the selected nations with a reliable system for forecasting crop production using remote sensing and related datasets. Crop inventory involves understanding the area occupied by a crop as well as its location, timing, and quantity[3]. To classify or categorize vegetation types, remote sensing datasets, such as Sentinel-2 and Landsat-8, were utilized, focusing on cloud-free information. Among various satellite datasets, Sentinel-2 has attracted significant interest from researchers because of its finer spatial resolution, free availability, and applicability in vegetation type/area classification and yield forecasting[2][3].

Remote sensing techniques involve gathering data about an object without requiring the sensor to make direct or physical contact with it[4]. Information is transmitted from an object to a sensor through electromagnetic radiation. The sensors utilized for these measurements can be situated a few meters away, several kilometers away on an aircraft, or even hundreds of kilometers away on satellites[5]. Multispectral remote sensing images summarize the critical integrated spectral and spatial characteristics of objects[6] Vegetation indices (VI) derived from satellite imagery serve as indicators of vegetation presence and health. Numerous studies on remote sensing applications have demonstrated that VI can be effectively utilized in crop monitoring and vegetation phenology characterization [7]. Digital image processing of satellite data employs various mathematical indices and algorithms as tools for image analyses. These features are based on reflectance characteristics, and indices have been developed to highlight the significant attributes of an image. Several indices can be used to identify vegetated areas in remote sensing images. Among these, NDVI is widely recognized and frequently used[8]. The Normalized Difference Vegetation Index (NDVI) is one of the most widely used and implemented indices and is calculated as a normalized ratio between the red and near-infrared bands of multispectral information. They are widely used to detect vegetation cover and crop health[9]. Those managing field data collected from plots will likely need to summarize the raster-based data associated with those plots. For example, it is essential to determine the Normalized Difference Vegetation Index (NDVI), precipitation, and elevation for each plot or its surrounding area. Zonal Analysis in GIS and Remote Sensing involves the application of computational techniques within Geographic Information Systems (GIS) and remote sensing to summarize and analyze spatial data across selected geographic regions. Zonal statistics are important in numerous fields, such as environmental studies, precision agriculture, and urban development. Zonal operations are spatial methods predicated on geographic features or

zones that are well defined by specific datasets. These datasets may be represented in vector (polygon) or raster format [10].Technological advances have facilitated the integration of Geographic Information Systems (GIS) with remote sensing, making it easier to conduct more complex analyses. This integration allows for the computation of statistical values within designated areas, a process known as Zonal Statistics, which has become an essential tool for summarizing data characteristics within specific geographic boundaries[11]
This study aimed to employ geospatial tools to monitor and track various crop types and their locations. It focuses on identifying key crops, such as cotton, sugarcane, sweet lime, and water bodies, using a method known as Random Forest. The accuracy of this method was evaluated through standard tests. It also examined NDVI values on different dates to assess crop growth and health. Furthermore, the study utilized areal statistics to provide detailed information about crops in specific areas, including the maximum, minimum, mean, and range of values. The goal is to enhance crop monitoring, aid farming decisions, and support sustainable resource use.

2.Methodology

The methodological framework depicted in Figure 1 presents the comprehensive workflow used in this study. The process begins with data acquisition from the Harmonized Landsat and Sentinel (HLS) datasets, followed by essential preprocessing steps, such as atmospheric correction and layer stacking. Subsequently, the Normalized Difference Vegetation Index (NDVI) was calculated because it is one of the most widely used indices for assessing vegetation health, growth patterns, and stress conditions. Additionally, classification was performed using the Random Forest algorithm to categorize primary land cover and crop types in the study area. The classified outputs were validated through accuracy assessment to ensure their reliability. Finally, zonal statistics were applied to analyze class-wise distributions and spectral responses across different zones, facilitating a clear understanding of the spatial variability in crop and land cover patterns. This structured approach ensured consistency and scientific rigour throughout the analysis..

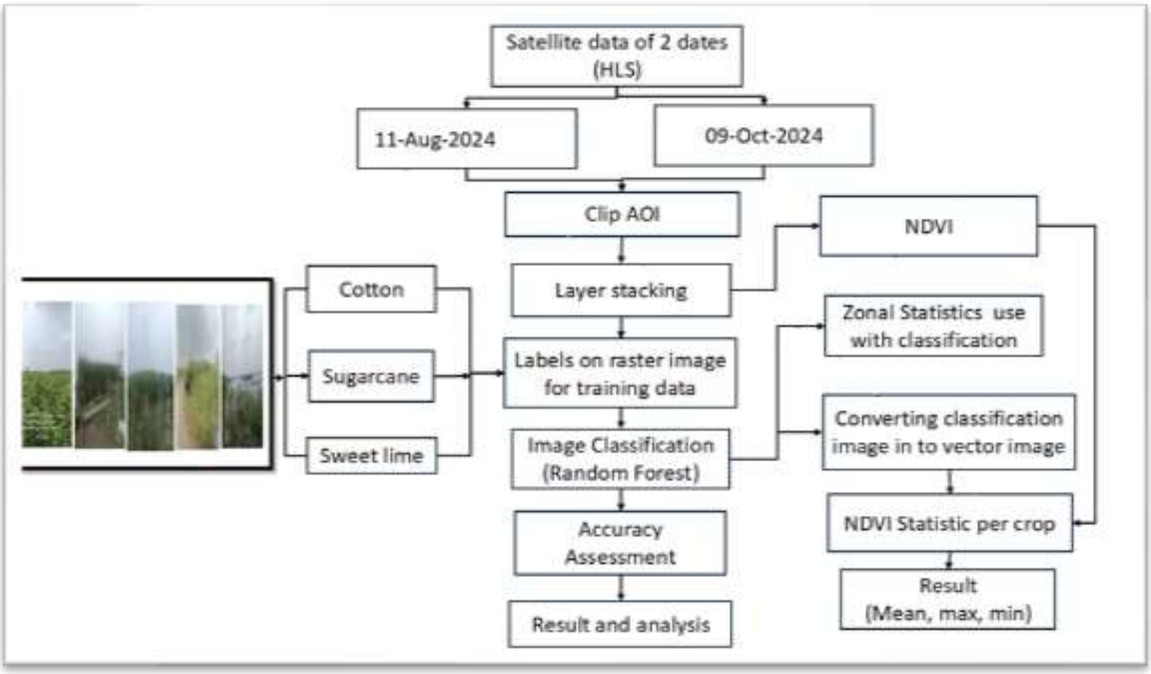

Fig1: Methodology adopted in the study (Prepared by the Shivani Bhosle)

2.1Study Area and Data Set

The study was supervised in MAHALAKSHUMIKHEDA village, Gangapur Taluka, Chh. Sambhajinagar, Maharashtra, India. The total topographical area is 527.65 hectares and the command area of the upper jaykwadi dam in Maharashtra. This area has a subtropical climate, with hot summers, moderate rainfall during the monsoon season, and mild winters. Farming in this area depends on irrigation from the dam, making it suitable for crop monitoring using multi-temporal imagery.

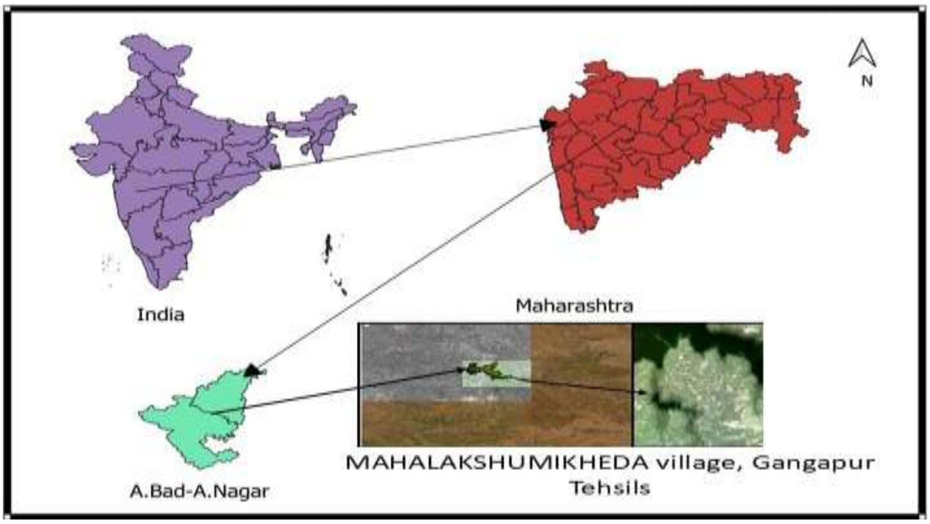

Fig.2: Study Area

2.1.2 Data Source and Band Selection

The key data source in this study, Harmonized Landsat & Sentinel-2 (HLS) data, were collected from the USGS. The Harmonized Landsat and Sentinel-2 (HLS) dataset is the product of a collaboration between NASA and the United States Geological Survey (USGS), which integrates data from the Landsat 8 Operational Land Imager (OLI) and the Sentinel-2 Multispectral Instrument (MSI)[12]. HLS data offers consistent 30-meter spatial resolution across all spectral bands, with observations available every 2-3 days for most of the Earth's land surface[13].The selected multi-temporal images were chosen for their capability to offer frequent observations without cloud interference. The Sentinel-2 component of HLS is mainly useful for cultivated studies because of its high spectral resolution, which includes bands specifically designed for vegetation analysis, such as the red-edge and near-infrared bands[14]. We employed HLS data from 2 different dates for this investigation.11 August 2024, 09. October 2024 The HLS dataset contained 13 spectral bands. These bands cover light from the visible to shortwave infrared (SWIR) spectrum. Six bands were selected: B2, B3, B4, B8A, B11, and B12. These bands are instrumental in agriculture for studying vegetation, identifying crops, and detecting water and soil stresses [15].

Table 1 shows these bands and their significance.

Band	Name	Wavelength (µm)	Application in Study
B2	Blue	0.45 – 0.51	This study identifies chlorophyll absorption, water bodies, and early-stage crop vigor.
B3	Green	0.53 – 0.59	Responsive to plant vigor and biomass, this metric is valuable for the monitoring of vegetation.
B4	Red	0.64 – 0.67	Key for chlorophyll absorption and crop stress detection
B8A	Narrow NIR	0.85 – 0.88	Assesses vegetation canopy, biomass, and NDVI calculation
B11	SWIR 1	1.57 – 1.65	Identifies vegetation moisture, soil water content, and drought stress
B12	SWIR 2	2.11 – 2.29	Differentiates vegetation types, detects senescence, and monitors water stress

2.1.3Ground Truth data

The quality and amount of field data affect the accuracy and reliability of crop maps when supervised machine learning is used [16]. In this study, we gathered field data from 100 different locations in the region to cover all types of terrains and features. We used GPS devices and smartphone apps to record the exact locations with longitude and latitude, along with photographs. To accurately geo-tag this location, as illustrated in Fig.3. Each of the 100 locations was chosen to include different land types, such as farms, towns, water areas, and other key features. At each location, we created a strong dataset with high-quality photos labelled with exact GPS locations. This helped to capture the overall view and ensured the spatial accuracy required for future analysis and classification.

Fig.3: Ground truth data from Mahalakshumikheda

3. Experimental approach

This study followed a systematic experimental workflow for crop identification and classification using two-date HLS satellite data. The workflow included data pre-processing, NDVI calculation, classification, Zon wise statistic and accuracy assessment

3.1Pre-processing and Preparation

Preprocessing is a vital step in remote sensing data studies, particularly when working with multitemporal satellite imagery. HLS data were used, which had undergone significant spectral and atmospheric correction. The HLS dataset provides spectrally corrected surface reflectance values from both Landsat and Sentinel-2 sources[17], which consider atmospheric elements, such as fine particles, clouds, and water vapor, thereby producing standardized reflectance values that are ready for analysis without requiring further modification. Layer stacking is a technique used in remote sensing and image processing to combine multiple spectral bands or images into a single multi-band composite[18]. Next, the multitemporal image was clipped to match the area of interest (AOI) representing the study area as shown in fig 2

Fig.3: Layer Stack Image

Images collected at two separate times were used to extract bands from the Red, Green, Blue, Near-Infrared, and Shortwave Infrared regions. These bands were combined into a single multiband raster consisting of 12 layers.

$$stack_{Image} = Stack(B_1^{t1}, B_2^{t1}, \dots B_{\dot{n}}^{t1}, \qquad B_1^{t2}, B_2^{t2} \dots B_n^{t2})$$

Where $(B_1^{t1}, B_2^{t1}, \dots . B_n^{t2}, B_2^{t2}{}_2, \dots ., B_n^{t2}$,The selected bands from two different dates were characterized, and Stack is a function used for band compositing that combines several spectral bands into one multispectral raster dataset. To prepare the ground truth data for training the machine learning models. To analyze raster image data for the MAHALAKSHUMIKHEDA village study using QGIS, import the 12-layer stacked raster file and create a false color composite. Digitalize training samples for different land cover classes and extract spectral values. Supervised classification is performed using the Semi-Automatic Classification Plugin, followed by post-classification processing. This process enabled crop monitoring and vegetation health assessment in the study area. In the framework of the study area, which relies on irrigation from the upper Jaykwadi Dam, image stacking of HLS data could provide valuable insights into crop development patterns and water resource management throughout the subtropical climate's distinct seasons.

3.2 NDVI Calculation

The normalized difference vegetation index (NDVI) takes advantage of how vegetation soaks and reflects light in the red and near-infrared regions of the electromagnetic spectrum[3].The Normalized Difference Vegetation Index (NDVI) is a widely used remote-sensing indicator for assessing vegetation health and density. It utilizes the differential absorption and reflection of light through vegetation in the red and near-infrared (NIR) areas of the electromagnetic spectrum. Healthy vegetation absorbs the maximum observable light for photosynthesis, while reflecting a large portion of NIR light[19]. We calculated all multi-temporal HLS images using surface reflectance values from the red (band 4) and NIR (Narrow NIR)(band 8A) combinations in QGIS according to the following formula:

$NDVI_{HLS} = \frac{B8A-B4}{B8A+B4}$ .....(1)

Where,

B4 (Red, 0.64–0.67 µm) † represents chlorophyll absorption.

B8A (Narrow NIR, 0.85–0.88 µm) † represents vegetation reflectance.

The NDVI index resultant in values ranging from -1 to +1. Higher positive values indicate denser, healthier vegetation, whereas lower values suggest sparse or stressed vegetation. It is a prevalent tool for mapping vegetation health, and is effective for monitoring desertification, drought evaluation, net main production of vegetation,crop growing conditions, etc.[3].

3.3Image classification

In crop monitoring, image classification involves categorizing images into predetermined classes using supervised classification techniques. In this study, the Random Forest classification technique was used for multitemporal images to classify crop types.

Random Forest is an aggregated machine-learning method that unites multiple decision trees for classification and regression tasks. It uses majority voting and feature randomization to create diverse trees, improve accuracy, and reducing overfitting[20]. This method handles multidimensional data, captures nonlinear relationships, and provides feature significance measures. Random forests are widely used in various fields owing to their versatility and effectiveness[21]. The formula for the Random Forest method is as follows:

$y^{\Lambda} = majority_{vote(T\,1\,(x),T\,2\,(x),\dots,T\,K\,(x))}$ .....(2)

y^{Λ} where Final predicted class label,x: Input feature vector, $T1(x), T2(x) \dots, TK(x)$ Predictions from each individual decision tree in the forest , majority vote Chooses the most frequent class label predicted by all trees. The following fig3. shows the classified image of random forest method

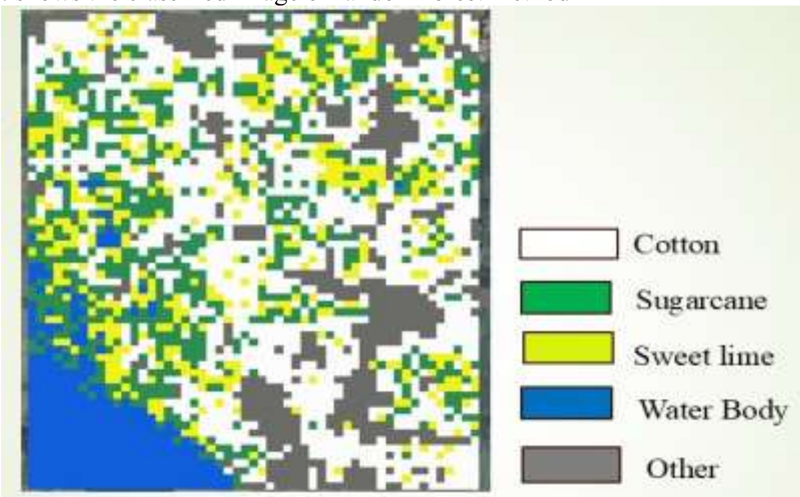

Fig3: Random Forst Classified Image

3.4Zonal statistics

Zonal statistics is a core technique used in Geographic Information Systems (GIS) to analyze raster data in relation to a specific geographic area. This method involves overlaying a raster dataset, which represents the variable of interest, onto a vector dataset that defines an area. Zonal statistics is a geospatial analysis technique used to calculate summary statistics for defined areas or zones within a raster dataset(such as mean, median, sum, minimum, maximum, and standard deviation)in the context of NDVI[22].

3.4.1Conversion of Raster to vector image

The alteration of a classified raster into a vector format is a step in spatial analysis when it is necessary to represent raster-based classification results as vector polygons[23]. This is typically done by the "Raster to Vector" tool, in QGIS which trace edges of each raster class and creates equivalent vector polygons for each single value. When multi-temporal images are classified using the Random Forest method, the result is a raster, where each pixel is assigned a value corresponding to a specific class: cotton = 1, sugarcane = 2, sweet lime = 3, water body = 4, and other = 5. To enhance the usability of these data for purposes such as area analysis, labelling, or map creation, they should be transformed into a vector layer.Following fig 4 shows the classified raster image into vector layer.

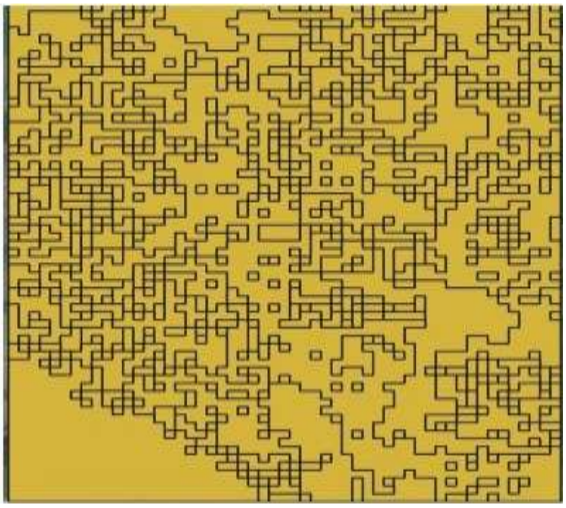

Fig4 : vector image of classified image

3.4.2Crop wise Statistical analysis with using Multidate NDVI map

The subsequent task involves calculating statistical values such as the minimum, maximum, mean, range, and sum for each zone (crop) using multidate NDVI maps (dated August 11, 2024, and October 09, 2024) along with a vector layer from a Random Forest classified image. This was achieved using the zonal statistics tool in the QGIS. In this research, the zones are derived from a classified raster image and include categories such as cotton, sugarcane, sweet lime, and waterbody.

The Mean of NDVI map calculated by following formula

$Mean_{z_i} = \frac{1}{n_i}\sum_{j=1}^{ni} R_i j$.....(3)

Where z_i denotes the ith zone ,R_{ij} the raster value of the jth cell within zone z_i ,n_i the total number of raster cell in zone z_i
The maximum value of NDVI map calculated by following formula
$Max_{zi} = \max(R_{i1}, R_{i2} \dots R_{in})$…(4)
whereMax_{zi} is the maximum value with in zone i, $R_{i1}, R_{i2} \dots R_{in}$ all raster cell values within zone i
The Minimum value of NDVI map calculated by following formula
$Min_{zi} = \min(R_{i1}, R_{i2} \dots R_{in})$…(5)
Where Min_{zi} is the minimum value with in zone I , $R_{i1}, R_{i2} \dots R_{in}$ all raster cell values within zone i
The Sum of NDVI map calculated by following formula
$sum_{zi} \sum_{i=1}^{n_i} R_{ij}$…(6)
Where sum_{zi} Sum of raster values for zone i , R_{ij} the value of the raster cell j within zone i
n_i total number of raster cells in zone i.
The Range of NDVI map calculated by following formula
$Range_{zi} = Max_{zi} - Min_{zi}$….(7)
Where z_i the zone, Max_{zi} maximum value with in zone, Min_{zi} minimum value with in the zone

4 Result

This section presents the results from NDVI based zonal statistic across two dates and accuracy by using Radom forest method. The index was assessed for its effectiveness in distinguishing major crop classes—cotton, sugarcane, and sweet lime—as well as water and other land-use categories

4.1 Zone-wise NDVI Variation and Class Statistics Across Two Dates

The zonal statistics table provides valuable information on the NDVI on August 11, 2024. Differences among the various crop types. Class 1, which corresponds to cotton (121 pixels), exhibited a high average NDVI of 0.64, indicating strong crop vitality. Class 2, associated with sugarcane (229 pixels), also demonstrated a significant mean NDVI of 0.61, which is consistent with its dense vegetation and significant biomass. Class 3, representing sweet lime (225 pixels), had a mean NDVI of 0.63, confirming stable crop health and productivity. On the other hand, Class 4, which refers to water bodies (21 pixels), recorded the lowest mean NDVI (0.40), as predicted by minimal vegetation reflectivity. Class 5, categorized under other land uses (74 pixels), presents a moderate NDVI (0.57), reflecting a mix of surfaces, such as barren land, settlements, or fallow fields. This analysis highlights that areas with cotton, sugarcane, and sweet lime show strong vegetation indicators, whereas water bodies and non-vegetated areas are clearly distinguished, thereby affirming NDVI as an effective tool for differentiating crops and land cover.

Table 2: NDVI-based zonal statistics of date 11-Aug-2024

	DN	count	unique	min	max	range	sum	mean
1	1	121	121	0.22346040606...	0.78086560964...	0.55740520358...	77.5380969585...	0.64081071866...
2	2	229	229	0.15090090036...	0.78435224294...	0.63345134258...	140.992528059...	0.61568789545...
3	3	225	225	0.15822784602...	0.79840618371...	0.64017833769...	142.445685518...	0.63309193563...
4	4	21	21	0.16772823035...	0.72596806287...	0.55823983252...	8.60020131869...	0.40953339612...
5	5	74	74	0.26235353946...	0.70898294448...	0.44662940502...	42.3955654365...	0.57291304644...

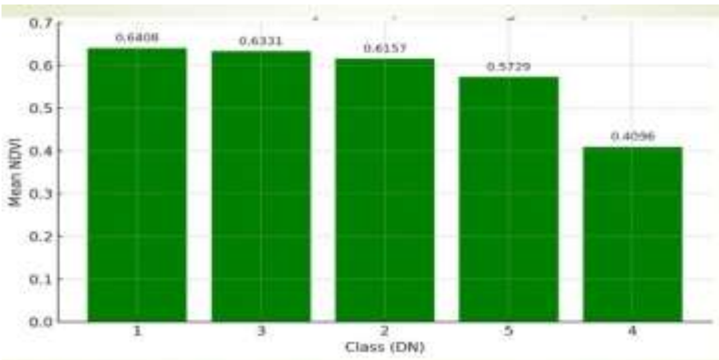

***Figure 5**. The mean NDVI values showed discrete variations across DN classes. DN class 1 recorded the highest NDVI (0.6408), followed by DN class 3 (0.6331), and DN class 2 (0.6157), indicating healthy vegetation vigor. DN class 5 displayed a moderate NDVI value (0.5729), whereas DN class 4 had the lowest NDVI value (0.4096), suggesting sparse or degraded vegetation cover.*

The table below presents the NDVI-based zonal statistics of date 09 -Oct-2024 that efficiently distinguish between vegetation classes and land cover types. Class 1 (cotton, 121 pixels) exhibited a high mean NDVI of 0.65, indicating healthy crop growth with a dense canopy. Class 2 (sugarcane, 229 pixels) showed a mean NDVI of 0.63, which is reliable because of its robust biomass and dense vegetation structure. Class 3 (sweet lime, 225 pixels) had a mean NDVI of 0.64, confirming a stable vegetation vigor and productive groves. In contrast, Class

4 (water bodies, 21 pixels) recorded the lowest mean NDVI of 0.41, as expected because of the low reflectance from the water surfaces. Class 5 (other land uses, 74 pixels) displayed a moderate mean NDVI of 0.60, representing mixed surfaces such as fallow land, barren patches, or settlements.

Table 3: NDVI-based zonal statistics of date 09 -Oct-2024

DN	count	unique	min	max	range	sum	mean
1	121	121	0.23224352300...	0.79600787162...	0.56376434862...	79.4627025160...	0.65671654971...
2	229	229	0.18259723484...	0.80667410790...	0.62407687306...	144.889774032...	0.63270643682...
3	225	225	0.20676402747...	0.81631368398...	0.60954965651...	146.142263526...	0.64952117122...
4	21	21	0.16330655578...	0.73499190807...	0.57168535228...	8.75789477032...	0.41704260811...
5	74	74	0.24769702553...	0.72733831405...	0.47964128851...	44.4323857984...	0.60043764592...

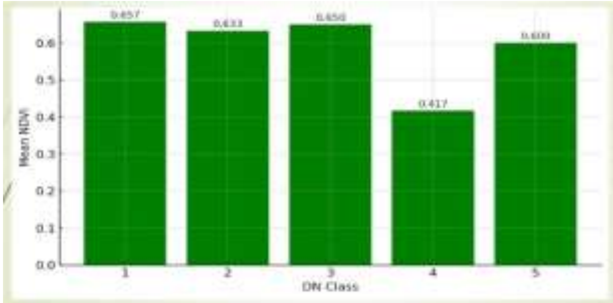

Figure 6. *The mean NDVI values differed among the DN classes. DN class 1 showed the highest mean NDVI at 0.657, followed by DN class 3 (0.650), and DN class 2 (0.633), all indicative of healthy vegetation vigor. DN class 5 presented a moderate NDVI value of 0.600, whereas DN class 4 had the lowest NDVI of 0.417, suggesting areas with water or no vegetation.*

4.2 Evaluate Classification accuracy(Random Forest)

The Random Forest (RF) algorithm achieved an overall classification accuracy of 81.17% and a kappa coefficient of 0.7539, reflecting a strong relationship between the RF-classified produce and reference data. Among the crop classes, cotton was classified as highly reliable (UA 73.07%, PA 90.36%), whereas sugarcane showed moderate accuracy (UA 69.01%, PA 67.96%). Sweet lime exhibited a high UA (88.57%) but relatively low PA (58.02%), indicating some omission errors. Water bodies were mapped with very high accuracy (UA 100%, PA 94.16%), whereas the "others" class displayed a high UA (100%) but low PA (34.16%) because of confusion with other land cover types. Overall, the Random Forest classification produced good results, particularly for the cotton, sweet lime, and water classes, although improvements are needed for sugarcane and the heterogeneous "other" category.

Table 4. Accuracy Evaluation of Random Forest

Crop Type	**UA (%)**	**PA (%)**
Cotton	73.07	90.36
Sugarcane	69.01	67.96
Sweet lime	88.57	58.02
Water	100.0	94.16
Others	100.0	34.16
Overall Accuracy	**81.17**	
Kappa	**0.7539**	

5.Conclusion

This study demonstrated that the integrating NDVI analysis, zonal statistics, and Random Forest method used for crop identification and agricultural monitoring. The results showed that NDVI provided a clear distinction between healthy and stressed vegetation, whereas zonal statistics helped quantify these variations at the regional level. The Random Forest classifier successfully differentiated crop types such as cotton, sugarcane, and sweet lime, along with water bodies and other land-cover classes, with satisfactory accuracy levels. These results confirm that combining spectral indices with machine learning offers a reliable and cost-effective method for precision agriculture.

Looking ahead, this approach can be further strengthened by incorporating multi-temporal datasets to monitor crop growth stages throughout the season. The use of higher-resolution satellite imagery or integration with UAV/drone-based data could improve the classification accuracy, especially for small or fragmented fields. Additionally, applying other advanced machine learning algorithms and comparing their performance with that of a Random Forest may provide deeper insights. This methodology also has potential applications in yield prediction, irrigation planning, and climate impact assessment. Overall, this study provides a strong foundation for scalable, data-driven agricultural management and can guide policymakers and farmers toward sustainable resource use.

Acknowledgements
The Authors would like to acknowledge the SARTHI Fellowship; CSMNRF 2022/2022-23/2341 for financial assistance and Dr. G. Y. Pathrikar college of Computer Science and Information Technology, MGM University, Chhatrapati Sambhajinagar for providing resources for the study

References

1. Rakholia, R., Tailor, J., Prajapati, M., Shah, M., Saini, J.R.: Emerging technology adoption for sustainable agriculture in India– a pilot study. J. Agric. Food Res. 17, 101238 (2024).
2. Singh, G., Singh, S., Sethi, G., Sood, V.: Deep Learning in the Mapping of Agricultural Land Use Using Sentinel-2 Satellite Data. Geographies. 2, 691–700 (2022).
3. Patel, J.H., Oza, M.P.: Deriving crop calendar using NDVI time-series. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XL–8, 869–873 (2014).
4. Zhang, J., Wang, P., Yao, F., Xu, Y., Guo, W., Li, L., Yang, L.: Advances in estimation methods of vegetation water content based on optical remote sensing techniques. Sci. China Technol. Sci. 53, 1159–1167 (2010).
5. Getahun, S., Kefale, H., Gelaye, Y.: Application of Precision Agriculture Technologies for Sustainable Crop Production and Environmental Sustainability: A Systematic Review. ScientificWorldJournal. 2024, (2024).
6. Karaburun, A., others: Estimation of C factor for soil erosion modeling using NDVI in Buyukcekmece watershed. Ozean J. Appl. Sci. 3, 77–85 (2010).
7. Maimaitijiang, M., Daloye, A.M., Erkbol, H., Sidike, P., Fritschi, F.B., Sagan, V.: Crop Monitoring Using Satellite/UAV Data Fusion and Machine Learning. Remote Sens. 12, 1357 (2020).
8. Xue, J., Su, B.: Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. J. Sensors. 2017, 1–17 (2017).
9. Kwan, C., Bernabe, S., Plaza, A., Ayhan, B., Li, J., Gribben, D.: An Accurate Vegetation and Non-Vegetation Differentiation Approach Based on Land Cover Classification. Remote Sens. 12, 3880 (2020).
10. Garcia, A.R., Filipe, S.B., Fernandes, C., Estevão, C., Ramos, G.: Zonal Statitic, https://atlas.co/glossary/zonal-statistics/, last accessed 2025/08/14.
11. Eltoum Masaad, M. (2025).: Zonal analysis in GIS and Remote sensing.
12. NASA Earthdata, https://earthdata.nasa.gov, last accessed 2024/12/12.
13. Berra, E.F., Breunig, F.M., Fontana, D.C., Yin, F.: Harmonized {Landsat} and {Sentinel}-2 {Data} with {Google} {Earth} {Engine}. Remote Sens. 16, 2695 (2024).
14. Saini, R., Ghosh, S.K.: CROP CLASSIFICATION ON SINGLE DATE SENTINEL-2 IMAGERY USING RANDOM FOREST AND SUPPOR VECTOR MACHINE. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLII–5, 683–688 (2018).
15. Hong, C., Prishchepov, A. V., Jin, X., Han, B., Lin, J., Liu, J., Ren, J., Zhou, Y.: The role of harmonized Landsat Sentinel-2 (HLS) products to reveal multiple trajectories and determinants of cropland abandonment in subtropical mountainous areas. J. Environ. Manage. 336, 117621 (2023).
16. Alami Machichi, M., mansouri, loubna El, imani, yasmina, Bourja, O., Lahlou, O., Zennayi, Y., Bourzeix, F., Hanadé Houmma, I., Hadria, R.: Crop mapping using supervised machine learning and deep learning: a systematic literature review. Int. J. Remote Sens. 44, 2717–2753 (2023).
17. Liu, Y., Hu, X., Gong, W., Gong, J.: Forest Type Identification with Random Forest Using Sentinel-1A, Sentinel-2A, Multi-Temporal Landsat-8 and DEM Data. Remote Sens. 10, 946 (2018).
18. Wu, H., Zhang, J., Lu, C., Zhao, S.: Remote Sensing Image Sharpening by Integrating Multispectral Image Super-Resolution and Convolutional Sparse Representation Fusion. IEEE Access. 7, 46562–46574 (2019).
19. Aini Rahmi, M., Parikesit, P., Withaningsih, S.: Vegetation change analysis using Normalized Difference Vegetation Index (NDVI) in Sumedang Regency. E3S Web Conf. 495, 2007 (2024).
20. Kulkarni, V.Y., Sinha, P.K.: Pruning of Random Forest classifiers: A survey and future directions. Presented at the June 30 (2012).
21. Ghosh, D., Cabrera, J.: Enriched Random Forest for High Dimensional Genomic Data. IEEE/ACM Trans. Comput. Biol. Bioinforma. 19, 2817–2828 (2022).
22. Esau, I., Miles, V. V, Kurchatova, A., Davy, R., Miles, M.W.: Trends in normalized difference vegetation index (NDVI) associated with urban development in northern West Siberia. Atmos. Chem. Phys. 16, 9563–9577 (2016).
23. Hamdani, Y., Xiao, G., Calvanese, D., Ding, L.: An Ontology-Based Framework for Geospatial Integration and Querying of Raster Data Cube Using Virtual Knowledge Graphs. ISPRS Int. J. Geo-Information. 12, 375 (2023).