Journal of Agrometeorology

ISSN : 0972-1665 (print), 2583-2980 (online)
Vol. No. 26 (4) : 505-508 (December - 2024)
https://doi.org/10.54386/jam.v26i4.2669
https://journal.agrimetassociation.org/index.php/jam

Short communication

Sugarcane acreage estimation using satellite imagery and machine learning

MEGHARANI B. MAYANI [1*] and RAJESHWARI L. ITAGI [2]

[1]*Government Polytechnic Belgaum, Electronics and Communication Engineering, Department of Collegiate and Technical Education, Government of Karnataka, and Visvesvaraya Technological University, Belagavi 590018, India*
[2]*KLE Institute of Technology, Electronics and Communication Department, Hubballi – Karnataka, and Visvesvaraya Technological University, Belagavi 590018, India*
[*]*Corresponding author email: megha.mayani@gmail.com*

Agriculture being the major farm practice in India due to 60% of land cover for cultivation. The recent climate change hazards are affecting this activity which is the backbone of Indian economy. As technology is moving at its pace and conquering over every sector in society in its own way, application in the agriculture sector is a significant contribution. Use of modern tools and technology is necessary to handle any shortfalls or surplus of the crops in a better way, by monitoring and predicting the crop yield before harvesting. Crop yield prediction is found essential for farmers (to plan on fertilizer input, water budgeting, pest management, etc.), industries (logistics, scheduling mill operations) and government (policy planning, decision making) for sustainable economy of the country (Everingham *et al.,* 2007). Geospatial technology and machine learning / deep learning can be a unique set of solutions to address the problems related to soil, water, land etc. (Altalak *et al.,* 2022). Satellite images with geospatial tools come in handy along with parallel computing facility aids in preparing accurate crop maps and yield predictions in which machine learning techniques play an effective approach.

India being the second largest sugarcane producer in the world, next to Brazil, sugarcane is termed as a multipurpose crop (Malik *et al.,* 2019) since it is used to make sugar, jaggery, khand sari, molasses, even paper. Karnataka state stands third to contribute sugarcane at national level. In Karnataka, sugarcane is grown as a *Kharif* crop and *Zaid*; come in varieties such as Eksali (12 months crop) and Adsali (18 months crop) to differentiate as a year crop and one-half year crop respectively. Every year, the state produces not less than 40 million tons of sugarcane. Being an essential commodity, a gap exists between the expected yield and actual yield taking the resources as constraints. The crop yield prediction requires the preliminary work of crop identification and machine learning/ deep learning algorithms have been explored and experimented by various researchers (Lonare *et al.,* 2022), to generate crop maps (crop identification). In the perspective of above-mentioned points in this research work, sugarcane crop is considered for crop identification and area estimation.

Study area

Belagavi district is popularly called as sugar bowl of Karnataka accounting for 35% of state sugarcane production (Sreedhar, *et al.,* 2022). The district is bounded by the north Karnataka at the foothills of the Sahyadri range (Western Ghats). The Zadshahpur village region of Belagavi district was selected for the proposed research work. The sugarcane farmland in Zadshahpur encompasses an area of 16 acres, lying between the latitude of 15^0 51' 1.30" N and longitude of 74^0 30' 16.81" E, at an altitude of 776 m above the mean sea level.

The growth phases of sugarcane crop (Eksali variety – annual crop) was identified as Germination (February- March), Tillering (April-June), Grand growth (July-September) and Maturity (October-December). Sugarcane varieties grown at the Zadshahpur village were mainly Co86032, CoC671, Co-09004 (Fig. 1). The diverse growth stages of the crop namely, varieties of green manure crops, ratoons etc. displayed a variation acquired by sensors which were collected through remote sensing satellite and depicted in images. Appropriate selection of time series satellite data was done to minimize the spectral overlap among different croplands (Luo *et al.,* 2021) and correspondingly increased the crop identification accuracy (Aghababaei *et al.,* 2021).

Article info - DOI: *https://doi.org/10.54386/jam.v26i4.2669*
Received: 21 July 2024; Accepted: 18 September 2024; Published online : 01 December 2024

Fig. 1: Sugarcane crop grown at Zadshahpur, Belagavi, Karnataka

Satellite data

Sentinel-2 image data from January to December for three years period (2021, 2022, 2023) were obtained from GEE public archieve data catalog. The dataset used in this study were having 10 m spatial resolution, 13 spectral bands and 5 days temporal resolution which are most suitable for small field size studies. With 72 images every year, for a period of 3 years, there were 216 images to process and analyze. But images with less than 0.01% cloud coverage were used for our research to get least corrupted sensor values that gave 12 clean images as data for processing. The Zadshahpur boundaries, train beds and test beds are uploaded as assets on the Google Earth Engine (GEE) cloud. Visible bands, NIR band, red edge band and SWIR bands were used as hyperparameters for training the machine learning algorithms. The parallel computing technology driven service of Google Earth Engine (GEE), was used as a platform to get remote sense data and to perform cloud computing with visualization analysis tools. (Luo *et al.,* 2021). The band values at surface reflectance level were used to create a mosaic of 12 images for each year. Hence GEE was used as a service for accessing multi sensor data along with computing. The NDVI was calculated as; NDVI = (B8 – B4) / (B8 + B4).

Gradient Tree Boost (GTB) for crop identification

Decision tree-based prediction has wide range of advantages (Arab *et al.,* 2021). Classification And Regression Tree (CART) are the simplest to implement, CART model involves selecting features. Random Forest (RF) is easy to use and flexible. An ensemble of uncorrelated trees is used in this algorithm. Gradient Tree Boost (GTB) machine learning algorithm can model very complex relational variables which is also an ensemble method. GTB includes averaging of target label, calculating residuals and then constructing a tree. This approach improved the learning process for our Zadshahpur in identifying sugarcane crop grown areas compared to CART and RF classifiers. Vegetation dynamics of sugarcane crop requires a gap free monitoring of vegetation over time, to discriminate it from other crops. So, ground truth data and supervised classification were used on time series NDVI data, to get sugarcane crop growth profile and differentiate it from other crops and fallow land. The same training data was fed to all three classifiers. After tuning hyperparameters GTB outperformed the other two for Zadshahpur. Temporal NDVI and SWIR were then given to all three classifiers. GTB gave better classification accuracy and then crop acreage estimation was predicted with GTB classifier.

Evaluation parameters in this research work, with remotely sensed image classification were accuracies, F1 score, Confusion matrix, kappa coefficient, etc. Overall accuracy and kappa coefficient were used as metrics for assessment. The training data was gathered by intermittent quarterly field visits to Zadshahpur for the years 2021, 2022, 2023 and previous years data was collected from reliable sources of past records from state government organization. To account for variability in climate, water content, soil, farming practice, etc., polygons were used from different farmlands. Because of this, at same timestamp, different growth stages of crops were used as input to algorithm. A total of 48 polygons (by doing data augmentation techniques) were used to train the machine learning models. Distinct testing polygons were selected to analyze the assessments effectively. Crop identification over small fields was carried out using machine learning algorithms and satellite image data. Remotely collected sensor values over a sample polygon used as a training site in Zadshahpur, for the years 2021, 2022, 2023 and growth profile indicator NDVI (Table 1).

The band values B3, B4, B5, B6, B8, B11, B12 corresponds to Green, Red, Red Edge1, RedEdge2, NIR, SWIR1, SWIR2 respectively. Green band (B3) with a scaling of 0.0001 over a spectral resolution of 10m has wavelength of 560nm for S2A and 559nm for S2B satellites. Spectral separability among the crops grown at the Zadshahpur was difficult since rice and sugarcane share similar spectral response over initial growing season, but time series analysis and unique transition timelines resulted in the covariance between them high and discrimination easy by identifying the sugarcane crop regions from others. Sugarcane crop maps were generated by CART, RF and GTB classifiers. Crop phenology-based inputs to the model became vital as farming practices varied region wise. Assessments were compared with and without these inputs using confusion matrix and knowing the misclassifications the overall accuracy and kappa coefficient were calculated and are presented in Table 2. Knowledge of sugarcane crop phenology that carry unique spectral signatures helped in our study to minimize the

Table 1: The calculated NDVI on observation dates during 2021-23

Observation date	NDVI	Observation date	NDVI	Observation date	NDVI
24-Jan-21	0.126	04-Jan-22	0.21	29-Jan-23	0.29
24-Jan-21	0.117	19-Jan-22	0.18	18-Feb-23	0.108
08-Feb-21	0.108	29-Jan-22	0.21	20-Mar-23	0.114
13-Feb-21	0.116	03-Feb-22	0.216	19-Apr-23	0.12
28-Feb-21	0.108	13-Feb-22	0.214	14-May-23	0.234
05-Mar-21	0.146	23-Feb-22	0.165	29-May-23	0.503
10-Mar-21	0.118	28-Feb-22	0.154	08-Jun-23	0.341
15-Mar-21	0.148	10-Mar-22	0.137	01-Sep-23	0.417
20-Mar-21	0.098	15-Mar-22	0.179	06-Oct-23	0.508
04-May-21	0.192	31-Oct-22	0.351	26-Oct-23	0.466
26-Oct-21	0.277	30-Nov-22	0.256	30-Nov-23	0.511
20-Dec-21	0.251	25-Dec-22	0.189	25-Dec-23	0.445

Table 2: Crop identification accuracy of classifiers without and with phenology factors

Classifier	Overall accuracy (in %)		Kappa coefficient	
	Without phenology factors	With phenology factors	Without phenology factors	With phenology factors
GTB	72.40	94.26	0.691	0.8931
RF	69.26	89.26	0.542	0.8412
CART	53.44	83.44	0.513	0.8123

Table 3: Acreage estimation of sugarcane by ML classifiers

Classifier	Acreage (in acres)	
	Sugarcane	Others
GTB	13.20	1.88
RF	12.50	1.78
CART	11.68	1.67

uncertainty in crop identification, were the parameters considered to train the same set of machine learning algorithms.

Improved results as overall accuracy to a greater extent is achieved for GTB from 72.4% to 94.26% (Table 2). With the application of machine learning algorithms, identification and crop map for sugarcane crop in the Zadshahpur, having sugarcane grown over 14 acres, rice grown over 0.5 acres, soyabean grown over 1acre and 0.5 acre with fallowland was selected for validation of the machine learning techniques used (after knowing the ground truth) was conducted and results are presented in Table 3. The pixels classified as *Sugarcane* and *Non Sugarcane* were quantified to get an approximate area of sugarcane grown croplands and non sugarcane grown regions i.e. others. From the results, it is evident that GTB performed better in predicting acreage estimation as it identified the sugarcane crop from others (rice, soya and fallow land) more accurately than other machine learning techniques.

ACKNOWLEDGEMENTS

Authors acknowledge the support in collecting ground truth data of crop, to faculty and staff of S. Nijalingappa Sugar Institute, Belagavi, Karnataka, India.

Source of funding: No source of funding.

Conflicts of Interests: The author(s) declare(s) that there is no conflict of interest.

Data availability: Could not share.

Author's contribution: **MB Mayani**: Writing original draft, Data analysis, Methodology Conceptualization, Data curation; **RL Itagi**: Investigating, Validation, Editing, Supervision, Reviewing.

Disclaimer: The contents, opinions, and views expressed in the research communication published in the Journal of Agrometeorology are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

Publisher's Note: The periodical remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

Aghababaei M., Ebrahimi A., Naghipour A, Asadi E. and Verrelst J. (2021). Vegetation Types Mapping Using Multi-Temporal Landsat Images in the Google Earth Engine Platform. *J.*

Rem. Sens., 13: 4683. https://doi.org/10.3390/rs13224683

Altalak M., Ammaduddin M., Alajmi A. and Rizg A. (2022). Smart Agriculture Applications Using Deep Learning Technologies: A Survey, *MDPI J. App. Sci.,* 12: 5919. https://doi.org/10.3390/app12125919

Arab S., Noguchi R., Matsushita S. and Ahamed T. (2021). Prediction of grape yields from time- series vegetation indices using satellite remote sensing and a machine-learning approach. *J. Rem. Sens. Apps. Soc. Env.,* 22: 100485, https://doi.org/10.1016/j.rsase.2021.100485

Everingham L., Lowe H., Donald A., Coomans H. and Markley J. (2007), Advanced satellite imagery to classify sugarcane crop characteristics, *Agron. Sustain. Dev.,* 27(2): 111-117. https://doi.org/10.1051/agro:2006034

Lonare A., Maheshwari B. and Chinnasamy P. (2022). Village level identification of sugarcane in Sangali, Maharashtra using open-source data, *J. Agrometeorol.,* 24(3):249-254. https://doi.org/10.54386/jam.v24i3.1688

Luo C., Qi B., Huanjun Liu, Guo D., Lu L., Fu Q. and Shao Y. (2021). Using Time Series Sentinel-1 Images for Object-Oriented Crop Classification in Google Earth Engine, *J. Rem. Sens.,* 13:561. https://doi.org/10.3390/rs13040561

Malik K., Verma D., Srivastava S., Mehta S., Kumari N., Verma A. and Singh P. (2019). Sugarcane Production and Its Utilization as a Biofuel in India: Status, Perspectives, and Current Policy, *Sugarcane Biofuels*, 123-138. https://doi.org/10.1007/978-3-030-18597-8_6

Sreedhar R., Varshney A. and Madhu D. (2022). Sugarcane crop classification using time series analysis of optical and SAR sentinel images: a deep learning approach, *Article Rem. Sens. Letters.,* 13(8): 812-821. http://dx.doi.org/10.1080/2150704X.2022.2088254