RESEARCH ARTICLE OPEN ACCESS

Machine Learning Approach to identify Sugarcane Crop grown area with Satellite Imagery using Soft Computing.

Megharani B Mayani*, Dr. Rajeshwari L Itagi**

**Senior Scale Lecturer (Electronics & Communication Engineering Department), Government Polytechnic Belgaum and Visvesvaraya Technological University – Belagavi 590018, India*

*** Professor (Electronics & Communication Engineering Department), K L E Institute of Technology Belgaum and Visvesvaraya Technological University – Belagavi 590018, India*

ABSTRACT
Prediction of a crop through satellite imagery is one of the applications of remote sensing technique. A variability in field over a period of years and location impose to develop a tool for field change analysis. Creating a crop yield map will be a solution providing a robust and accurate prediction system designed to work with small field boundaries as an input. The real challenge is with a single pixel per plot being used to analyze and predict the crop, health status of crops and yield. Sugarcane being identified as an essential commodity; it can be identified with acceptable accuracies using good spatial satellite imageries. The five stages of sugarcane cycle will be Germination, Leaf development, Tillering, Stem Elongation and Harvestable parts development. Datasets comprising temporal data over the crop calendar is a part of training the model. To address the issue a cost-effective solution to the growers can be a boon. A user interface can be done to calculate cane yield, commercial cane sugar and sugar yield after an extensive testing. Forecast of the yield in mid of the season will guide the farmer on fertilizer input (which is mostly mishandled) and of course policy makers. Decision Tree based machine learning models were used to analyze; Extreme Gradient Boost has shown better results in comparison with others. Acceptable accuracy with minimum training data and for a small field area is observed with 70/30 as train/test data sets.
Keywords - confusion matrix, decision tree, extreme gradient boost classifier, indices, sentinel, yield map.

Date of Submission: 02-04-2024 Date of acceptance: 13-04-2024

I. INTRODUCTION

India, the land of agriculture; it's not just for Indians but our mother land feeds many people across the globe through its different variety of grains, spices, fruits, vegetables etc. but unfortunately days are not far for us to face food crisis no matter what the reason would be but the coming days are for sustainability [1]. Hence a need to combat the above by revolutions in farming and also the preparedness for the coming uncertainties is the need of hour. By the use of modern tools and technology a farmer can not only just monitor but also can predict the yield before harvesting so that shortfall and surplus of the crops can be handled in a better way [2]. Technology is moving at its pace and conquering over every sector in society in its own way; application of it in agriculture sector also will definitely not disappoint us. Every farmer is struggling with a different cause [3] for poor yield which is demotivating many of them and stop farming. Replacement of traditional methods by simple ease ways (with the use of technology) fusion with the experience and an expert involvement will be a way to strengthen agriculture in our country [4]. Geospatial technology and Machine learning/ Deep learning can be a unique set of solution to address many problems related soil, water, land etc. [5].

India is the 2nd largest sugarcane producer in the world next to Brazil. Sugarcane is termed as a multipurpose crop since it is used to make sugar, jaggery, khandsari, molasses, even paper [6]. Also, sugarcane is a distinguished cash crop which is called as a water guzzling crop as well. Karnataka stands third to contribute sugarcane at national level after Uttar Pradesh and Maharashtra. Every year state produces not less than 40 million tons of sugarcane. Belagavi is popularly called as Sugar Bowl of Karnataka accounting for 35% of state sugarcane production [7]. As sugarcane has been declared as an essential commodity a gap exists between the expected yield and actual yield taking the resources as constraints. Any initiative to help farmers to ease farming and also educate them on increasing their yield will be encouraging. With this as an objective an attempt is made to identify sugarcane crop grown area and its approximate yield. In Karnataka sugarcane is a kharif crop and come in varieties such as Eksali and Adsali to differentiate as a year crop and one-half year crop respectively [8]. With Ten taluks and 21 working sugar industries Belagavi feel itself proud in producing more than one crore ton of sugarcane every year by not only benefitting the state in revenue but also central excise department [9].

Section II details the methodology used to carry out research, different stages, study area selected, satellite data used, cloud computing technique used. Section III discusses the decision tree algorithms training, with split data, results obtained with each tree and assessments with validation.

II. WORK METHODOLOGY

A precise estimate of sugarcane grown area in different geographical zones is an essential input for policy makers. Currently the distribution of sugarcane grown is not accurately quantified, despite their social and economic importance. The primary aim of the work was to study the potential use of geospatial technology for mapping of sugarcane grown area for developing proof of concept which will not only help the Department of Agriculture and Farmers but also the sugarcane industries. A baseline database can be created to map and monitor the crop growth and also implementing it across the country will help to take up remedial measures for better yields by scaling up the learnings from this study and which will in turn help to meet the sustainability goal [10].

This pilot study is envisaged for the application of geospatial technology for assessment of sugarcane grown area in the selected study areas using Sentinel Satellite (COPERNICUS) imageries and Machine Learning Techniques [11]. The major objectives of the study are as follows:

1. Generation of sugarcane grown map of the study area and its spatial extent.
2. Estimation of the yield.
3. Validation of the above from beneficiaries.

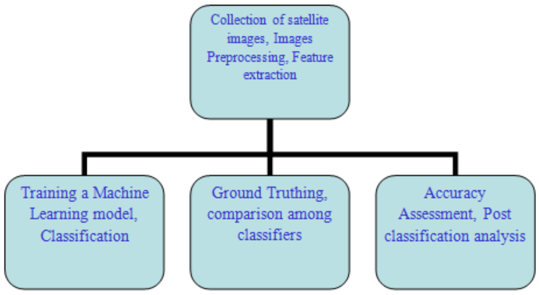


Fig. 1 : Workflow of the research performed

The study area was identified in consultation with the local experts, industrialists and government statistics. These areas not only represent good sugarcane grown in the state, but also have unique post farming products in the state. This necessitates the development and implementation of state-of-the-art machine learning-based models to suit the microcosms of farmer needs across the country [12]. Success of the pilot study in this region will be a way forward for seamless upscaling of the same to state and national level crop grown assessment systems.

The district encompasses an area of 13414 sq. kms lying between the latitude of 15.852792N and longitude of 74.498703E. The district lies approximately about 776m above the mean sea level. The district is bounded on the North Karnataka at the foothills of the Sahyadri range (Western Ghats). Figure 1 shows the work flow used to carry the research in the area of interest. Initially a small area of around 16 acres with mixed cropping in the region of interest was considered to work upon (Figure 2). The ground truth was collected by field visits (current) and from a reliable source (past records from state government organization). Google Earth Engine service is used as a platform to get remote sense data and also perform cloud computing with visualization analysis tools. The public data catalogue is used for study. Plenty of datasets are available for scientific and research surveys. More than around previous 30 years of data is available for access. The Earth Engine API is available in Python and JavaScript, making it easy to harness the power of Google's cloud for our own geospatial analysis

[13]. The web-based code editor is used for fast, interactive algorithm development with instant access to petabytes of data. Among various datasets Sentinel 2 images during latter half of year 2021 with surface reflectance were used [14]. Specific key bands and GNDVI (Green Normalized Difference Vegetation Index) are used as inputs for training the model [15]. Cloud coverage is nullified by creating mask which masks pixel with null values.

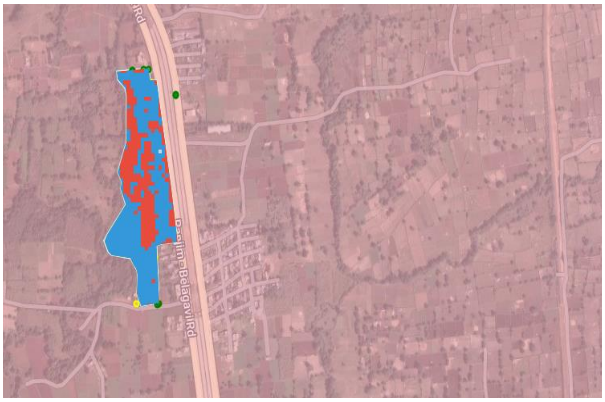
[Courtesy: Google Earth Engine]
Fig. 2 : Study Area used for research

III. RESULTS AND DISCUSSION

A small region of interest chosen in Belagavi district of Karnataka state was applied supervised classification with machine learning algorithms. The machine learning algorithms which were applied on area of interest were Random Forest, CART and Extreme Gradient Boost classifiers. The algorithms were used with default arguments and the results obtained are tabulated below in table No.1 with accuracy figures. The classified maps are as shown below in figures. As evident from the maps shown in Fig. 3, of classified images, the Extreme Gradient Boost classifier outperformed the other type classifiers. The confusion matrix was used as metric to get the error estimate. The number of mis classifications or wrong predictions / False Negatives were found to be comparatively less in Extreme Gradient Boost classification.

Table 1: Accuracies of ML classifiers used for study.

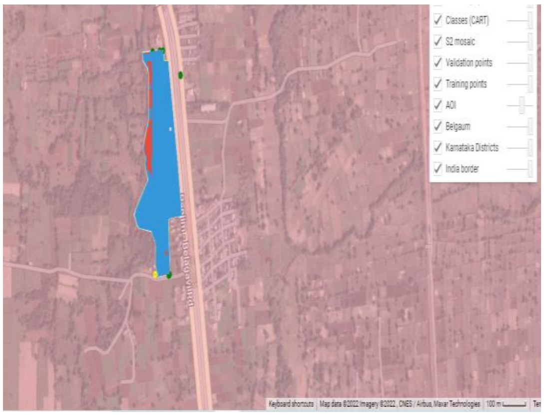
Sl.No.	Machine Learning Algorithm	Training Accuracy	Overall Accuracy
1	Extreme Gradient Boost Classifier	100%	92.4%
2	Random Forest Classifier	100%	88.1%
3	Classification And Regression Tree	100%	79.6%



[Courtesy: Google Earth Engine]
Fig. 3a : Extreme Gradient Boost classifier map results

A small set of training points gave a convincible result though. The chart below depicts the area grown with sugarcane crop and other crops. The actual sugarcane crop grown area is 14.5 acres and remaining with other crops. The proposed model depicts 14.8 acres with sugarcane grown (Fig.4). Results are guiding to carry forward and prepare a yield map to regional or national scale to forecast the yield very much before the actual harvest time. This not only aids growers but also mill operators.

[Courtesy: Google Earth Engine]
Fig. 3b : Random Forest classifier map results



[Courtesy: Google Earth Engine]
Fig. 3c : CART classifier map results

Since the sugarcane growing season mostly passes along the monsoon cloud coverage of pixel and masking those may take away some crucial deciding features; fusion of Synthetic Aperture Radar (SAR) data can be a rescue to be experimented yet. Exploration of other indices like Leaf Area Index (LAI), Green Chlorophyll Index (GCI) can be appended to sharpen the inputs given to the model to train it in a better way. Since single point images (peak growth time) ground data was used for training an alternate option would be multi point ground data as input.

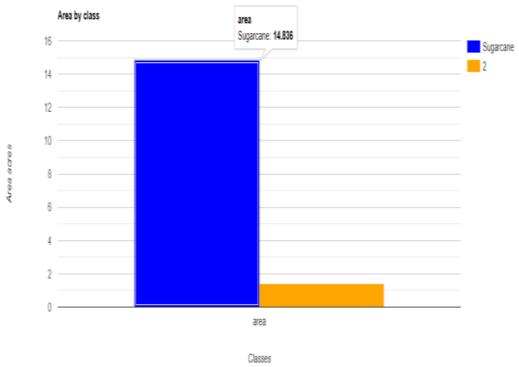


Fig. 4 : Extreme Gradient Boost classified chart showing crop grown area.

Future scope can be exploration of Deep learning models and high resolution data to train the model. Any small improvisation in the model accuracy can still be a approachable way to get the optimized results at the end user. Integration of wireless sensor network, automation and learning models with a user interface will be a reality in coming years taking farming at ease and precise. Though the other data (weather, climate, metereology etc) need to be connected to produce a full fledge support to the farmer through a stand alone application interface. An user interface with support for crop growers right from selection of crop, decision on crop calendar, fertilizer/ pesticide input, yield forecast, expert chatbots, etc. can greatly be a boon to farmers.

Acknowledgements

We acknowledge meaningful discussions with agronomists of S Nijalingappa Sugar Institute, Belagavi for educating us about sugarcane crop and its breeds, cultivation methods and its uses.

REFERENCES

[1]. Article in COMMUNITY-BASED SUSTAINABLE LIVELIHOODS, “WATER-EFFICIENT SUGARCANE FARMING”.

[2]. Food and Agriculture Organization of the United Nations, National Remote Sensing Centre (NRSC), Indian Space Research Organization (ISRO) and Department of Space, Govt. of India project report on “Spatial Inventory of Agroforestry Resources using Geospatial Technology and AI with Very High-Resolution Images” 2022.

[3]. J. P. Clemente, G. Fontanelli, G. G. Ovando1, Y. L. B. Roa, A. Lapini and E. Sant, “GOOGLE EARTH ENGINE: APPLICATION OF ALGORITHMS FOR REMOTE SENSING OF CROPS IN TUSCANY (ITALY)”, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-3/W12-2020, IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS 2020), 22–26 March 2020, Santiago, Chile.

[4]. Girish A Chavadappanavar, “Impact of COVID-19 on Sugarcane Growers and Sugar Industry-Situational Assessment in Karnataka State”, International Journal of Science and Research (IJSR) ISSN: 2319-7064 SJIF (2020): 7.803, Volume 10 Issue 9, Sep 2021.

[5]. Prabhu kadadi and Dr.Jahanara, “Extent of Adoption of Organic Sugarcane Cultivation Practices in Belgaum District of Karnataka State”, Volume 3, Issue 5, May – 2018 International Journal of Innovative Science and Research Technology.

[6]. Maha Altalak, Mohammad Ammaduddin, Amal Alajmi and Alwaseemah Rizg, “Smart Agriculture Applications Using Deep

Learning Technologies: A Survey", MDPI journal in Applied Science 2022, 12, 5919.
[7]. James H Cock, "Centre International for Agriculture Tropical Sugarcane growth and development", Article in International Sugar Journal · November 2003.
[8]. Rahul Sreedhar, Avnish Varshney and Dhanya Madhu, "Sugarcane crop classification using time series analysis of optical and SAR sentinel images: a deep learning approach" Article in Remote Sensing Letters - June 2022, Vol. 13, No. 8, 812–821.
[9]. Megharani B. Mayani and Dr. Rajeshwari Itagi, "Machine Learning Techniques in Land Cover Classification using Remote Sensing Data", 2021 International Conference on Intelligent Technologies (CONIT) Karnataka, India. June 25-27, 2021.
[10]. Teerapat Butkhot and Pipat Reungsang, "Assessment of machine learning on sugarcane classification using Landsat-8 and Sentinel-2 satellite imagery", Asia-Pacific Journal of Science and Technology: Volume: 26. Issue: 04. Article ID.: APST-26-04-08.
[11]. Sara Tokhi Arab,Ryozo Noguchi, Shusuke Matsushita and Tofael Ahamed, "Prediction of grape yields from time-series vegetation indices using satellite remote sensing and a machine-learning approach", Journal of Remote Sensing Applications: Society and Environment 22 (2021) 100485.
[12]. Escobar González José Luis, "REMOTE SENSING FOR CROPS IDENTIFICATION", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-3/W6, 2019 ISPRS-GEOGLAM-ISRS Joint Int. Workshop on "Earth Observations for Agricultural Monitoring", 18–20 February 2019, New Delhi, India.
[13]. J. P. Clemente, G. Fontanelli, G. G. Ovando1, Y. L. B. Roa, A. Lapini and E. Sant, "GOOGLE EARTH ENGINE: APPLICATION OF ALGORITHMS FOR REMOTE SENSING OF CROPS IN TUSCANY (ITALY)", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-3/W12-2020, IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS 2020), 22–26 March 2020, Santiago, Chile
[14]. Chong Luo, Beisong Qi, Huanjun Liu, Dong Guo, Lvping Lu, Qiang Fu and Yiqun Shao, "Using Time Series Sentinel-1 Images for Object-Oriented Crop Classification in Google Earth Engine", An article in Journal of Remote Sensing, 2021 Vol 13 Page 561.
[15]. Md Yeasin, Dipanwita Haldar, Suresh Kumar, Ranjit Kumar Paul and Sonaka Ghosh, "Machine Learning Techniques for Phenology Assessment of Sugarcane Using Conjunctive SAR and Optical Data", Remote Sensing journal. 2022, 14, 3249.
[16]. Tracy Whelen and Paul Siqueira, "Coefficient of variation for use in crop area classification across multiple climates", International Journal of Application Earth Observation Geoinformation, 67 (2018) 114–122.