



TEMPORAL PROCESSING OF REMOTELY SENSED OPTICAL IMAGES OF A SUGARCANE FARMLAND ON GOOGLE EARTH ENGINE

¹Megharani B Mayani, ²Dr. Rajeshwari L Itagi

¹Senior Scale Lecturer (Electronics & Communication Engineering Department), ²Professor (Electronics & Communication Engineering Department)

¹Government Polytechnic Belgaum, Electronics and Communication Engineering – Department of Collegiate and Technical Education – Government of Karnataka, and Visvesvaraya Technological University – Belagavi 590018, India

Abstract: Remote sensing is gaining popularity in agriculture systems [1] due to the advancements in technology which are heading towards precise spatial and temporal data sensed by satellites carrying the minute significant information. Despite considerable maturity of Earth Observation satellites in land and water monitoring; atmospheric attenuation has a big role to play in Optical satellites. Snow / cloud cover / aerosols degrade the recorded values considerably. Agriculture based applications such as crop identification, crop health, pest management, yield prediction, etc. are done by analyzing satellite images with huge data involved; but the indices used for above mentioned applications can be accurately computed if and only if the sensor values are reliable. Cloudy pixel value makes the analysis less accurate. To handle such pixel values a technique to reconstruct or predict the missing values is much needed. Also, the Machine learning / Deep learning models [2] have proven to be efficient when fed with a gap free data. Research has been done to find a technique to get the best estimate of the missing pixel values. Rolling statistics (Moving Average), Savitzky Golay filtering, Interpolation with Savitzky Golay filtering was done on Normalized Difference Vegetation Index (NDVI) over a period of four years at a sugarcane farmland region. Using Google Earth Engine [3] as a cloud computing service and satellite dataset provider the NDVI values obtained with above techniques were compared with the actual ground values. The observations and inferences are discussed in this paper. Research conducted showed that interpolation fused with filtering gave better approximations of missing values compared to other statistical methods.

Index Terms - Normalized Difference Vegetation Index (NDVI), Cloud Coverage, Google Earth Engine (GEE), Moving Average (MA), Savitzky Golay (SG) filter, Interpolation.

I. INTRODUCTION

Remotely sensed satellite images play a major role in the field of Agriculture to identify crop type, monitor crop health status, infected crop and its effect on yield, etc. The sensors carried onboard by Earth Observation Satellites downlink valuable information which is analyzed at ground stations to infer significant observations. Over decades of years optical remote sensing has revolutionized the approach in environmental monitoring, disaster monitoring, climate change, etc. Multiple spectral instruments (have few sensors) and Hyperspectral instruments (have hundreds of sensors) falling in the range of 300nm – 2400nm of electromagnetic spectrum records the radiation reflected from the Earth's surface. But clouds/ aerosols forming an interface between the Earth's surface and satellite sensors modulate the values and as a result the recorded values suffer from noise [4]. Cloud cover poses a serious challenge in data analysis as they degrade the values to a greater extent. Since time series data carries significant information required to detect changes of land use and land cover [5] holds a special place in Agriculture based applications too. The missing observations (gaps) in the time series due to cloud coverage / snow or any atmospheric attenuation have limited the development of time series applications. This indeed has triggered research to reconstruct these missing values so that a gap free time series data can be generated to feed a Machine Learning / Deep Learning model [6] to optimize the analysis.

Normalized Difference Vegetation Index (NDVI) is one of the important indices [7] to quantify the greenness of the vegetation which also clues about vegetation density and plant health status. It is calculated using Red band and Near Infrared band (NIR) values as given below:

$$NDVI = (NIR - Red) / (NIR + Red)$$

This NDVI value range from -1.0 to +1.0. Where negative values indicate water / snow [<0.0], zero nearing positive values indicate bare soil [$0.0 - 0.1$], high positive values indicate sparse vegetation [$0.1 - 0.4$] and higher positive values indicate dense vegetation [>0.5]. NDVI time series is proven to be effective in depicting the vegetation phenology and vegetation activity identification [8]. A gap free temporal NDVI composite is an important input to the learning models to perform more accurately in extrapolation predictions. *R Agarwal, et al* have suggested use of Machine learning for estimating NDVI of cloudy pixels, *Florian Gerber, et al* suggested Quantile regression to predict missing values in Spatio – temporal remote sensing data, *Xinkai Liu, et al* suggested envelope detection based Savitzky Golay filter as a better method to reconstruct NDVI time series. Here an attempt to apply various techniques on small sized farm (few acres) with sugarcane crop grown as *Eksali* (annual crop) [9,10] was done to find the best method to get a gap free NDVI temporal composite of which the details and findings are discussed as below.

II. MATERIALS AND METHODS

Google Earth Engine service [11] is used as a platform to get remote sense data and to perform cloud computing with visualization analysis tools. One of the paradigms – changing features of Earth Engine is the ability to access decades of imagery without the previous limitation of needing to download, organize, store and process this information. Plenty of datasets (around 300 Peta Bytes of data) are available for scientific and research surveys. More than around previous 30 years of data is available for access. The public data catalogue is used for study. 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 [12]. The web-based code editor is used for fast, interactive algorithm development with instant access to petabytes of data. Among various datasets Sentinel 2 level 1C images during years 2019, 2020, 2021 and 2022 with surface reflectance are used [13]. Visible Red band and Infrared band values are collected to compute NDVI. The observed NDVI mosaic over a period of four years over our study area (Fig. 1a & Fig. 1b) is depicted in Fig.2.

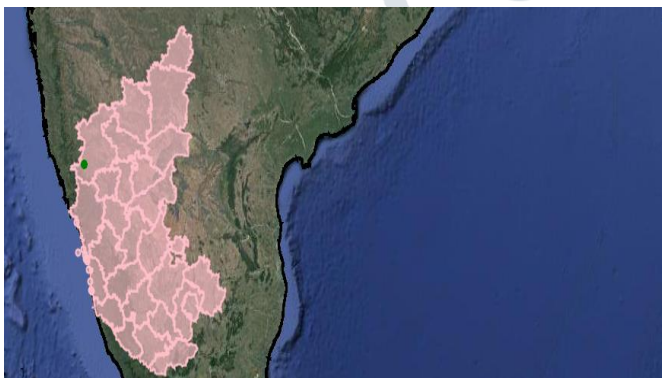


Fig. 1a: Study Area used for Research

[Courtesy: Google Earth Engine]

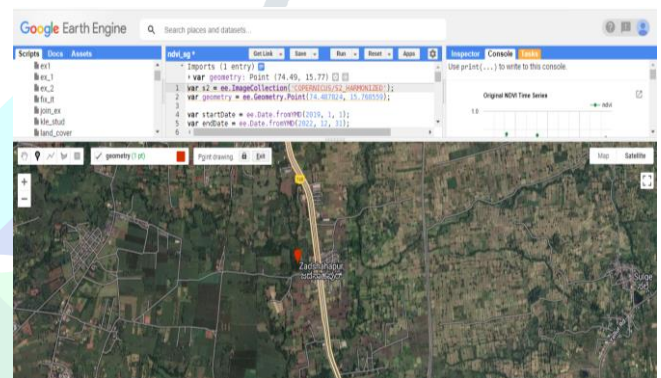


Fig. 1b: Region of Interest of Sugarcane farm

[Courtesy: Google Earth Engine]

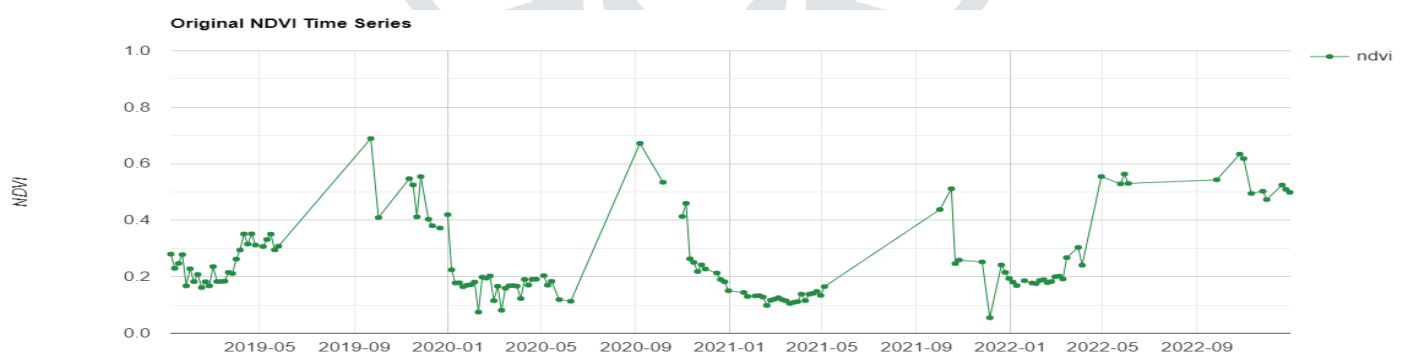


Fig. 2: NDVI plot over Region of Interest for consecutive four years: 2019 – 2022 [Courtesy: GEE]

Sentinel 2 (A & B) together give us data at 10m spatial resolution and 5days temporal resolution; same is used for our research. As evident from Fig. 2 the discontinuity [14] in NDVI profile and also the aberrations which is the effect of noise in the remotely sensed satellite data. The atmosphere which is an interface between the remote sensing instrument and target is the cause for it [15], though its unavoidable but to deal with it is critical as any learning method requires no gap data as input to give reliable accurate outputs. Hence a need to smoothen the temporal curve by optimizing the approximates. The methodology adopted to reconstruct the NDVI values is shown in Fig. 3.

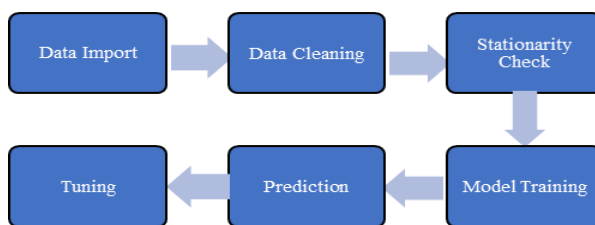


Fig. 3: Workflow of the research performed

Data is imported from GEE dataset, not a number (Nan) are dropped in data cleaning process. Then check on stationary is done followed by model training i.e., rolling statistics (Moving Average), Time Interpolation, Savitzky Golay Filtering are performed to predict the NDVI value. Assessment is done by finding Mean Absolute Error (MAE). Tuning of the parameters (time window size, order of filter, temporal neighbor size) was done to optimize on predictions.

Also, as evident from data extraction with reference to dates viz, Table 1, there are no images available with less than 20% cloud coverage during the monsoon season (June to September) in every year. Since sugarcane crop growth is very impulsive and dominant during monsoon and plant phenology during monsoon contains many vital information namely growth status, pests attack, stress condition information, etc. Hence crucial to perform preprocessing of this data before giving to any machine learning or deep learning model for further analysis or prediction. Cloudy pixel’s information on different bands was needed. Instead of depending on Synthetic Aperture Radar Satellite data [16]; combination of statistical methods was performed by collecting the temporal neighbor’s values and filling the gaps. Though; the validation of same was done from the field measured values. Spectroradiometer (Spectra Vista Corporation’s portable instrument – SVC H512i data) is used in the field of region of interest to collect the exact spectral information that are also called as ground truths. Also, with the help of Global Positioning System (GPS) [Garmin OREGON 650] device and Quantum Geographic Information System (QGIS) – an open-source software was used to create shape files of selected small region of interest that can be uploaded as assets in GEE to perform further analysis.

Table 1: Sentinel 2 images dates and corresponding NDVI values extracted in GEE system

Year	Day/ Month	NDVI Values	Year	Day/ Month	NDVI Values	Year	Day/ Month	NDVI Values	Year	Day/ Month	NDVI Values
2019	05-Jan	0.28	2020	05-Jan	0.225	2021	19-Jan	0.144	2022	04-Jan	0.181
	10-Jan	0.23		10-Jan	0.178		24-Jan	0.13		09-Jan	0.169
	15-Jan	0.247		15-Jan	0.179		03-Feb	0.132		14-Jan	0.175
	20-Jan	0.278		20-Jan	0.164		08-Feb	0.133		19-Jan	0.186
	25-Jan	0.168		25-Jan	0.17		13-Feb	0.128		29-Jan	0.178
	30-Jan	0.228		30-Jan	0.172		18-Feb	0.098		03-Feb	0.176
	04-Feb	0.183		04-Feb	0.181		23-Feb	0.117		08-Feb	0.187
	09-Feb	0.208		09-Feb	0.075		28-Feb	0.121		13-Feb	0.19
	14-Feb	0.162		14-Feb	0.199		05-Mar	0.126		18-Feb	0.179
	19-Feb	0.182		19-Feb	0.195		10-Mar	0.119		23-Feb	
	24-Feb	0.168		24-Feb	0.203		15-Mar	0.115		28-Feb	0.183
	01-Mar	0.236		29-Feb	0.115		20-Mar	0.105		05-Mar	0.2
	06-Mar	0.183		05-Mar	0.166		25-Mar	0.11		10-Mar	0.202
	11-Mar	0.183		10-Mar	0.081		30-Mar	0.112		15-Mar	0.192
	16-Mar	0.184		15-Mar	0.159		04-Apr	0.139		30-Mar	0.268
	21-Mar	0.215		20-Mar	0.168		09-Apr	0.115		04-Apr	0.304
	26-Mar	0.211		25-Mar	0.169		14-Apr	0.138		29-Apr	0.241
	31-Mar	0.262		30-Mar	0.167		19-Apr	0.141		24-May	0.555
	05-Apr	0.295		04-Apr	0.123		24-Apr	0.148		29-May	0.529
	10-Apr	0.352		09-Apr	0.191		29-Apr	0.134		03-Jun	0.563
15-Apr	0.316	14-Apr	0.17	04-May	0.165	26-Sep	0.53				
20-Apr	0.352	19-Apr	0.191	01-Oct	0.438	26-Oct	0.543				
25-Apr	0.312	24-Apr	0.191	16-Oct	0.511	31-Oct	0.634				
05-May	0.307	29-Apr		21-Oct	0.247	10-Nov	0.618				
10-May	0.332	04-May	0.204	26-Oct	0.259	25-Nov	0.495				
15-May	0.351	09-May	0.17	25-Nov	0.252	30-Nov	0.503				

20-May	0.295	14-May	0.184	05-Dec	0.055	20-Dec	0.473
25-May	0.308	24-May	0.119	20-Dec	0.242	25-Dec	0.525
22-Sep	0.689	08-Jun	0.113	25-Dec	0.215	30-Dec	0.509
02-Oct	0.409	06-Sep	0.672	30-Dec	0.194		0.499
11-Nov	0.547	06-Oct	0.534				
16-Nov	0.525	26-Oct					
21-Nov	0.412	31-Oct	0.413				
26-Nov	0.555	05-Nov	0.46				
06-Dec	0.404	10-Nov	0.263				
11-Dec	0.381	15-Nov	0.25				
21-Dec	0.372	20-Nov	0.219				
26-Dec		25-Nov	0.242				
31-Dec	0.42	30-Nov	0.227				
		15-Dec	0.213				
		20-Dec	0.19				
		25-Dec	0.182				
		30-Dec	0.151				

Using Image Collection to Image Collection Joins in GEE, Moving average was performed to get the missing values. The obtained temporal graph of NDVI after Moving Average method is shown in Fig. 4.

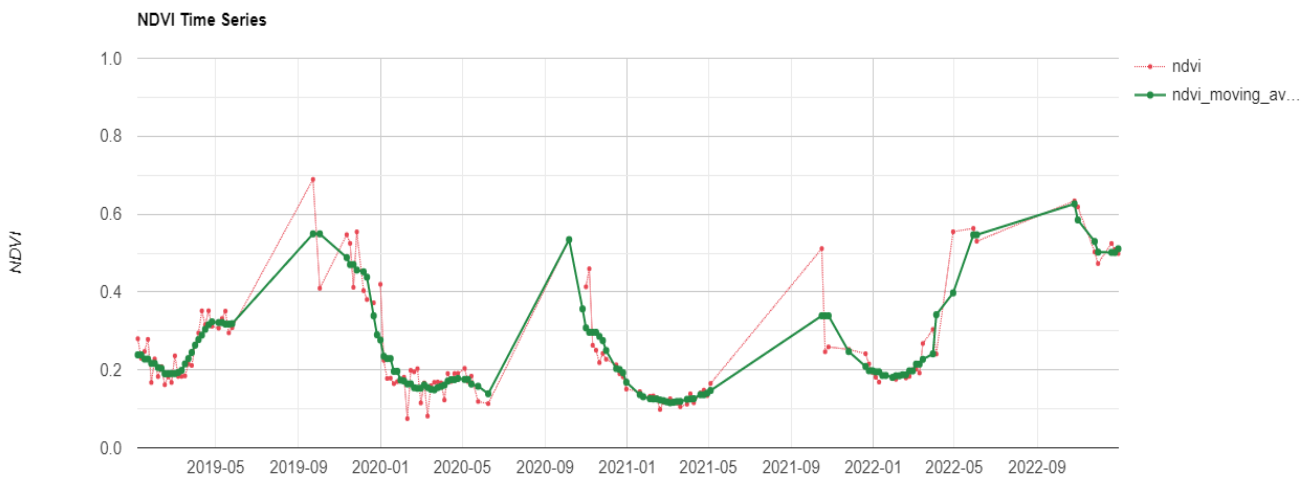


Fig. 4: NDVI plot over Region of Interest with Moving Average [Courtesy: GEE]

The moving average is calculated for each individual missing NDVI value using expression

$$Y [i] = \{1/M\} \sum_{j=0}^{M-1} X [i + j]; \quad \text{where } Y: \text{NDVI value, } X: \text{time stamp, } j: \text{window size}$$

The sharp rises and falls in NDVI values are smoothed by Moving average but long gaps were unattended. Creating a regularly spaced time series [17] and using temporal neighbors NDVI to interpolate the missing values and finally applying the SG filter gave the best results after hyper tuning the parameters. NDVI composite after combining time series interpolation and Savitzky Golay filtering is shown in Fig. 5. These values were found close to ground truth values after validation. The interpolation used was as below;

$$Y = Y1 + (Y2-Y1) * ((t - t1) / (t2 - t1))$$

Where Y = interpolated image,
 Y1 = before image, Y2 = after image
 t = interpolation timestamp
 t1 = before image timestamp
 t2 = after image timestamp

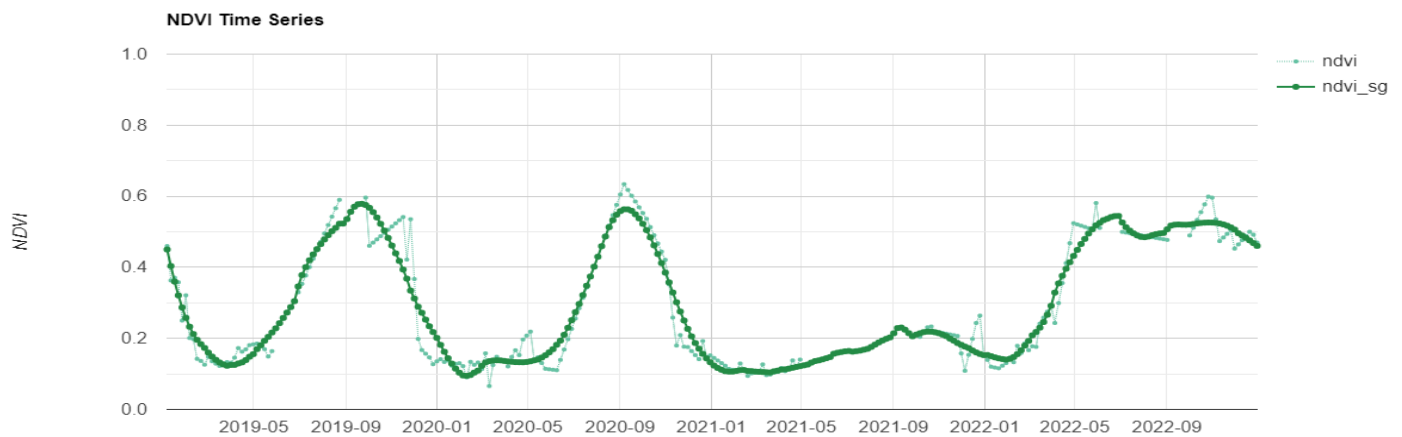


Fig. 5: NDVI plot over Region of Interest with combining time Interpolation and SG Filter
[Courtesy: GEE]

III. RESULTS AND DISCUSSIONS

The NDVI temporal profile after filling the missing pixel values by rolling statistics and curve fitting was performed on the same region of interest over same time stamp keeping other parameters same gave different observations with a Mean Absolute Error of 1.813 and 1.402 respectively for Moving Average and Interpolated SG Filter. It is very convincing that combining interpolation and filtering with proper tuning of parameters are relatively closer to real field values. These results can guide to carry forward and prepare a yield map [18] based on temporal NDVI pattern to regional or national scale to forecast the yield very much before the actual harvest time. 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 with optical data [19] can be a rescue to be experimented yet. Exploration of other indices like Enhanced Vegetation Index (EVI), Leaf Area Index (LAI), Green Chlorophyll Index (GCI) on deep learning models [20] can also be used to compute with observed optical sensor values to check with accuracy for above studied methods. A generic robust technique can be obtained for any number of variants. The success of the pilot study in this region will be a way forward for seamless up scaling of the same to other applications of time series processing.

REFERENCES

- [1] 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", New Delhi, India. Feb. 2019.
- [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. Ovando, 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, Vol. XLII-3/W12-2020, IEEE Latin American GRSS & ISPRS Remote Sensing Conference Santiago, Chile, March 2020.
- [4] Alber Ipia, Gilberto Camara and et. al, "Remote sensing comparison of cloud cover detection algorithms on Sentinel -2 images of the Amazon Tropical Forest", Journal of Remote Sensing Vol 12, 19th April 2020, pp.1284.
- [5] James H Cock, "Centre International for Agriculture Tropical Sugarcane growth and development", Article in International Sugar Journal · Nov. 2003.
- [6] Maha Altalak, Mohammad Ammaduddin, Amal Alajmi and Alwaseemah Rizg, "Smart Agriculture Applications Using Deep Learning Technologies: A Survey", MDPI Journal in Applied Science, Vol. 12, 2022, pp. 5919.
- [7] Xinkai Liu, Lingyun Ji, et. al., "A method for reconstructing NDVI time – series based on envelope detection and Savitzky Golay filter", International Journal of Digital Earth 2022, Vol. 15, No. 553 – 584.
- [8] Yanli Chen, Liping Feng, Jianfei Mo, Weihua Mo, Meihua Ding, Zhiping Liu, "Identification of Sugarcane with NDVI Time Series Based on HJ-1 CCD and MODIS Fusion", Springer Journal of the Indian Society of Remote Sensing, Vol. 48, Issue 2, 2020, pp. 249–262.
- [9] Girish A Chavadappanavar, "Impact of COVID-19 on Sugarcane Growers and Sugar Industry-Situational Assessment in Karnataka State", International Jour. of Science and Research, Vol. 10, Issue 9, Sep 2021. ISSN: 2319-7064

- [10] Prabhu Kadadi and Jahanara, “Extent of Adoption of Organic Sugarcane Cultivation Practices in Belgaum District of Karnataka State”, *Int. Jour. of Innovative Science and Research Technology*, Vol. 3, Issue 5, 2018.
- [11] Megharani B. Mayani and Rajeshwari Itagi, “Machine Learning Techniques in Land Cover Classification using Remote Sensing Data”, *International Conference on Intelligent Technologies (CONIT) Karnataka, India*. June 2021.
- [12] Masoumeh Aghababaei, Ataollah Ebrahimi, Ali Asghar Naghipour, Esaail Asadi and Jochem Verrelst, “Vegetation Types Mapping Using Multi – Temporal Landsat Images in the Google Earth Engine Platform”, *Journal of Remote Sensing*, Vol. 13, 2021, pp 4683.
- [13] 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*: Vol. 26. Issue 04. Article ID.: APST-26-04-08.
- [14] Zhipeng Tang, Giuseppe Amatulli and et. al, “Spectral Temporal Information for Missing Data Reconstruction (STIMDR) of Landsat Reflectance Time Series”, *Journal of Remote Sensing* Vol 14, 2022, pp.172.
- [15] Florian Gerberan and et. al, “Predicting missing values in spatio – temporal remote sensing data”, *IEEE transactions on Geo Science and Remote sensing*, Dec 2017.
- [16] 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*. Vol.14, 2022, pp.3249 - 3267.
- [17] Subrina Tahsin, Stephen C M, et. al., “Optical Cloud Pixel Recovery via Machine Learning”, *Journal of remote sensing* 2017, Vol. 9, No. 527.
- [18] Shyamal S. Virnodkar, Vinod K. Pachghare, V. C. Patil and Sunil Kumar Jha, “Application of Machine Learning on Remote Sensing Data for Sugarcane Crop Classification: A Review”, *Book of ICT Analysis and Applications*, Springer, 2020, pp. 539-555.
- [19] Sara Tokhi Arab, Ryozi Noguchi, Shusuke Matsushita and Tofael Ahamed, “Prediction of grape yields from time-series vegetation indices using satellite remote sensing and a machine-learning approach”, *Jour. of Remote Sensing Applications: Society and Environment*, Vol. 22, 2021.
- [20] 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* - Vol. 13, No. 8, June 2022, pp.812–821.