Analysis of The Water Scarcity Using Time-Series Rainfall and Satellite Dataset

[1]Monali B. Jadhav, [2]Sandeep V. Gaikwad, [3]Prof. Ratnadeep. R. Deshmukh

Department of Computer Science and Information Technology, Dr Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India, 431004.

monalij28@gmail.com, sandeep.gaikwad22@gmail.com, rrdeshmukh.csit@bamu.ac.in

Abstract: Rainfall is the primary key driver of an ecosystem, and it is responsible for sustaining life on the planet. Conversely, the rainfall deficiency creates a severe problem in the region, like water shortage for domestic, industrial and agricultural sectors. Furthermore, the region's land cover severely affects the barren land increment and decrement of the vegetation cover. To estimate and analyze the affected area using a manual survey is a very complex and tedious task. Therefore, meteorological and satellite imagery data provides a solution for effectively assessing the Land Use Land Cover (LULC) of the area. In the present research study, the historical rainfall data of the past 39 years and Sentinel 2 satellite data of the year 2016 and 2019 of the Vaijapur tehsil were used for the analysis. The Standard Precipitation Index (SPI) algorithm was applied to the historical rainfall dataset to investigate the region's meteorological drought. Furthermore, the ground truth points were used for the generation of training and testing pixel for classification using the Random Forest (RF) classifier. The Overall Accuracy (OA) was obtained 92% and kappa 0.9 in 2016 and 88.69% OA and 0.85 Kappa in 2019.

IndexTerms **- LULC, Random Forest, SPI, meteorological drought, Atmospheric correction**

I. INTRODUCTION

Drought is a natural phenomenon that reoccurred in every climatic division of the world. Therefore, the millions of people's livelihood affects severely than the other disaster [1,2]. Drought is classified into the three categories such as meteorological, hydrological and agricultural drought. The erratic rainfall behaviors or rainfall deficiency lead to meteorological drought conditions [3]. Moreover, low rainfall leads to a reduction in soil moisture. Therefore, the crop cannot sustain good health due to reduced soil moisture in the field. Such type of phenomenon is known as agricultural drought [4,5]. The reduced rainfall is responsible for the streamflow and presence of the surface water. Such conditions create water scarcity for drinking and irrigation, leading to hydrological drought. The monitoring and assessment of the drought are essential for the planning and decision-making support to the government and non-government agencies [6,7,8,9].

The SPI was developed at Colorado state university, USA, in 1993, useful for the rainfall analysis in every climatic division of the world [10,11,12]. It provides the output at various time scales like the short term and long term by analyzing the historical rainfall dataset. The short-term interval is used to analyze the soil moisture. In contrast, the longer-term interval is used to analyze water scarcity by monitoring the water level in a river, reservoir, lake, and groundwater region [13]. Therefore, the SPI helps analyze the land cover and meteorological drought of the region [14,15].

Similarly, the satellite provides the periodical observation of the earth, which helps natural resource monitoring and change analysis. Various popular satellites provide periodical observation with the earth surface's sizeable historical dataset such as MODIS, AVHRR, Landsat, Spot, AWIFS, and Sentinel satellites [16,17,18]. Generally, the optical RS satellite provides an observation into various bands like VIS (Visual infrared), SWIR (Shortwave Infrared) regions of the electromagnetic spectrum. Therefore, it is significantly helpful for the LULC change analysis, which is essential for planning and decision making in various operations like settlement planning, disaster mitigation, agriculture, barren land estimation etc. [19,20,21,22]. It has beneficial applications for the agriculture sector like crop health monitoring, irrigation management, estimation of crop acreage, and yield [23,24,25,26,27]. The following section describes the conducted experiment in detail.

II. STUDY AREA

Figure 1 shows the study area of the Vaijapur, Aurangabad district, Maharashtra, India. The average rainfall is 520-750mm, and temperature ranges (34-42°C) maximum and the region's minimum temperature. The Aurangabad district is a part of Marathwada territory, which comes under the scanty rainfall region. Therefore, the entire region is highly vulnerable to the drought disaster. Moreover, such types of disasters are responsible for severe damage to the region's livelihood. The economy of the region is based on agricultural and associated businesses.

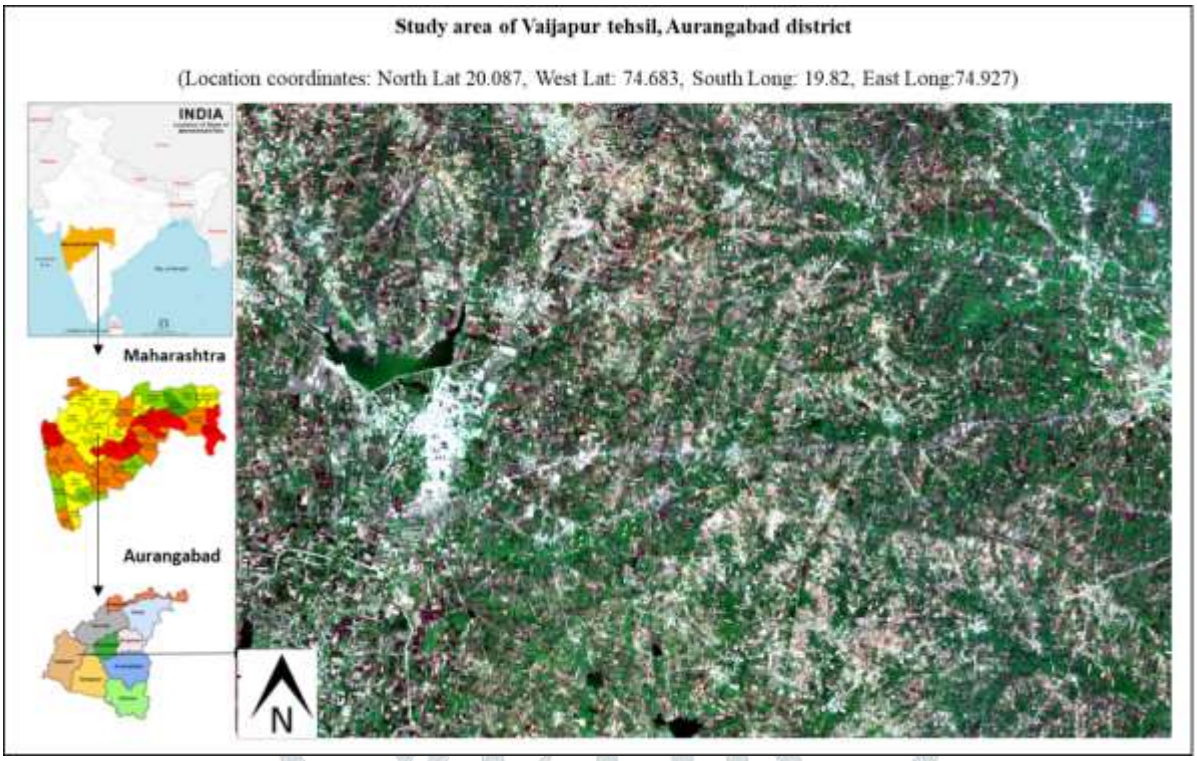

Figure 1. Study area of the Vaijapur tehsil

III. MATERIALS AND METHODS

Dataset

We have obtained the ground truth, rainfall and satellite dataset for the experimental analysis. The site visit was carried out on the day of the satellite overpass. The Ground Truth (GT) data was collected from the Samsung J7 smartphone. The GT data includes the GPS coordinates, photographs, and field description. Also, the rainfall data of the 39 years was obtained from the IMD (Indian Meteorological Department).

Furthermore, the Sentinel 2 satellite dataset was downloaded from the ESA (European Space Agency) scientific hub. Sentinel datasets dated 27/10/2018 and 25/10/2016 were downloaded from the ESA. The product was available in the standard level 1C, which is orthorectified and geometric corrected images. The Sentinel 2A image includes the total 13 bands observation, which includes 10m, 20m and 60m spatial resolutions.

Methodology

Figure 2. shows the methodology of the conducted experiment. We have used rainfall and sentinel satellite images and GT dataset for the water scarcity analysis. A total of 39 years (1981-2019) of monthly rainfall was used for the computation of the SPI index. The results were analyzed for the investigation of the meteorological droughts with the help of Sentinel satellite imagery. The following section shows the detailed information of the SPI computation.

Standardized Precipitation Index(SPI)

The Standardized Precipitation Index(SPI) is based on the gamma density function, which analyses the region's short-term and long-term meteorological drought episode. The SPI is a primary variable in the early drought monitoring system [28]. The THOM method [29] was used for the computation of the α and β parameters. In contrast, the αi indicates the yearly rainfall values. Moreover, the cumulative probability F(αi) at αi precipitation was computed using the gamma cumulative distribution function. Finally, the conversion of F(αi) to a standard normal random variable (SPI values) along with the calculated mean value to zero(0) and variance to one (1) was noticed. The SPI provided the range of +2 to -2, including the seven classes, which indicated the meteorological drought category. The positive values show the normal to wet condition, whereas negative values show the normal to extreme dry condition. Gamma density function is given [10,13,29] as follows:

$$f(x_i) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x_i^{\alpha-1} e^{-x_{i/\beta}} \quad (1)$$

Where β and α are the scale and shape parameter. The bigger the shape value is closer to the normal distribution curve.

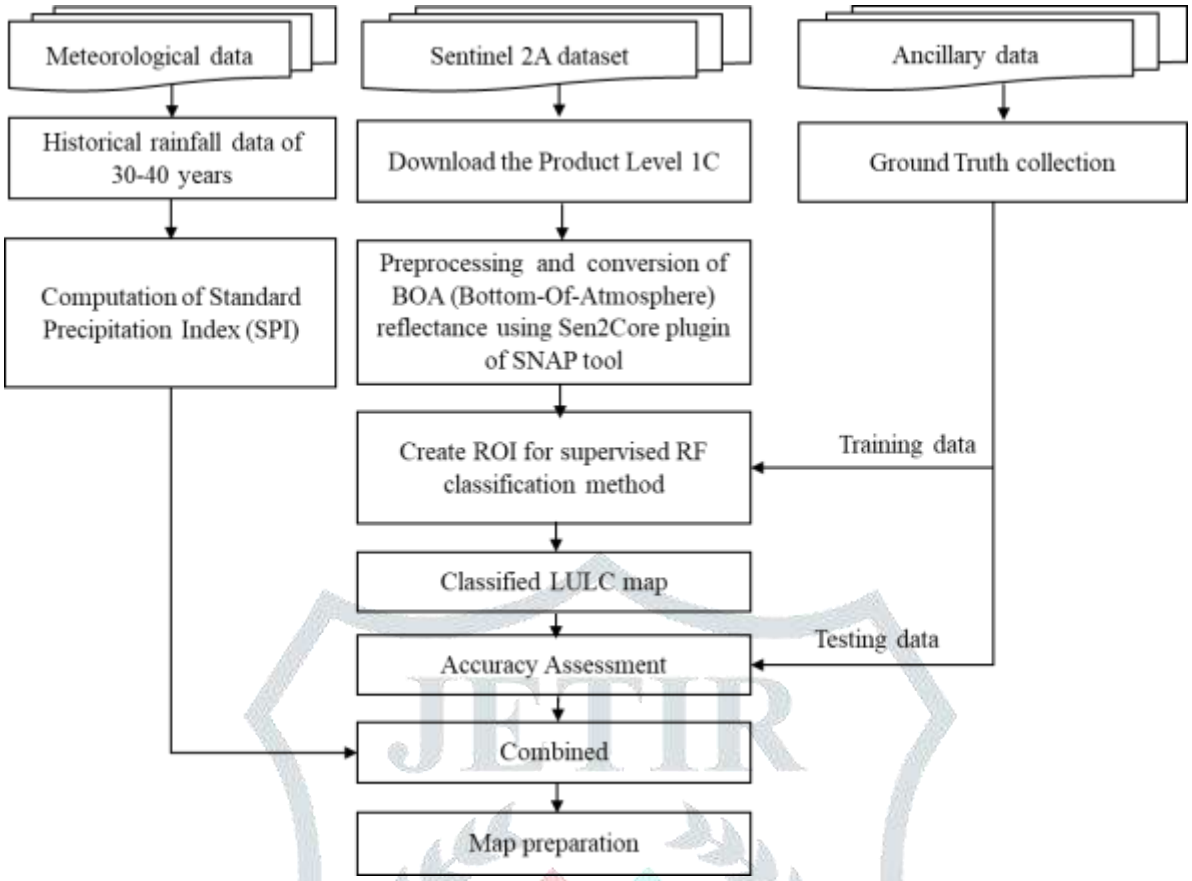

Figure. 2. Methodology used in the present research study

The x_(i)(>0) is the rainfall within I consecutive months, i.e. i time scales.

$$x_i^{(j)} = \sum_{k=1}^{i} P_{j,k,}\ j = 1,2,3 \ldots, n \quad (2)$$

Where P_(j,k,) is the prediction value of kth month of jth year. n is the number of years. We have used a 1-month time scale to calculate SPI, therefore the value of i=1,

The Gamma function Γ(α) is given as:

$$\Gamma(\alpha) = \int_0^{\infty} t^{\propto -1}\, e^{-t} dt \quad (3)$$

The α and β are the estimated with the approximation of Thom as follow:

$$\alpha = \frac{1}{4A}\left(1 + \sqrt{1 + \frac{4A}{3}}\right) \quad (4)$$

$$\beta = \frac{xi}{\alpha} \quad (5)$$

where $A = 1n(\bar{xi}) - \frac{1}{n}\sum_{j=1}^{n} 1n((x_i)_j)$

It is based on the probability density function Eq (1), the cumulative probability g(x_i) at the chosen time scale is given as follow:

$$g(x_i) = \int_0^{x_i} f(x_i)dx_i = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \int_0^{x_i} x_i^{\propto -1}\, e^{-x_{i/\beta}}\, dx_i \quad (6)$$

The probability of total no precipitation is given as below

$$F(x_i = 0) = \frac{m}{n} \quad (7)$$

Where m denotes the number of zero rainfall value in the computed data sequence whereas cumulative probability can be written as:

$$H(x_i) = F(x_i = 0) + \big(1 - F(x_i = 0)\big)g(x_i) \quad (8)$$

Finally, $H(x_i)$ can be transformed to SPI using the following equations by Milton.

$$SPI = -\left(t - \frac{C_0 + C_1 + c_2 t^2}{1 + d_1 t + d_1 t^2 + d_3 t^2}\right), t = \sqrt{1n\left(\frac{1}{H(x_i)^2}\right)}, for\ 0 < H(x_i) \leq\ 0.5 \quad (9)$$

$$SPI = -\left(t - \frac{C_0 + C_1 + c_2 t^2}{1 + d_1 t + d_1 t^2 + d_3 t^2}\right), t = \sqrt{1n\left(\frac{1}{H(x_i)^2}\right)}, for\ 0 < H(x_i) \leq\ 1 \quad (10)$$

Where c_(0)=2.515517,c_(1)=0.802853,c_(2)=0.010328,d_(1)=1.432788, d_(2)=0.189269 and d_(3)=0.001308.

The variable (SPI values) along with mean to zero and variance to one was noticed. More technical details regarding the SPI algorithm are given by Zhou [30]. SPI ranges between +2.0 to -2.0, including seven classes [30]. Classification of the region in different categories can be done based on values of SPI. Table 1 shows the classification criteria of meteorological drought levels based on SPI values.

Table 1. SPI criteria

SPI Range	Drought classes
2.0to +∞	Extremely wet
1.5 to 2.0	Very wet
1.0 to 1.5	Moderate wet
-1.0 to1.0	Near Normal
-1.5 to -1.0	Moderate drought
-2.0 to -1.5	Severe drought
-∞ to -2.0	Extreme drought

Preprocessing of Satellite images

The preprocessing of the satellite imagery is essential before further classification. The ESA provides the geometrically corrected sentinel satellite imagery known as level 1C product. It also provides the software tool known as SNAP (Sentinel Application Platform) for preprocessing and spatial analysis operations. The preprocessing operation includes the atmospheric, radiometric and geometric corrections of the Level 1C product images. The sen2core plugin of the SNAP was used for the preprocessing of sentinel images. The TOA (Top-Of-Atmosphere) values convert into the BOA(Bottom-Of-Atmosphere), known as the Level 2A product level of the sentinel. Furthermore, the supervised Random Forest (RF) classifier was chosen for the LULC map creation.

Use of Ground Truth (GT) for training and testing classifier

The RF required the ROI (Region of Interest) as input for the class generation for the classification. The ROI's were created for the training data with GT data collected at the time of the field visit. It helps to understand the features at the time of the training and testing the classifier. Finally, the classifier generated a region map, which is validated by using the testing GT data.

Supervised classification using Random Forest (RF) algorithm

Breiman has invented the classification and regression tree technique called Random Forest [31]. It creates the randomly and iteratively sample data and variables to generate a large group or known as the forest of classification trees. RF is based on ensemble learning techniques and built upon multiple Decision Trees(DCT). Each DCT is built using the original training dataset, and the remaining part of the training features validates it. RF works similar to the Maximum Likelihood classifier, where it creates the cluster based on similar spectral values. RF required the two essential parameters like (i)input file (ii)number of trees, (iii)the max number of samples per class (default value is 5000) (iv)vector training (v) feature band selection. The RF is based on the Gini index as an attribute selection measure, which measures an attribute's impurity to the classes. The Gini index is shown in equation 11. classes. The Gini index is shown in equation 11.

$$\sum_{j\neq i}\sum (f(C_i,\mathrm{T})/|\mathrm{T}|)(f(C_j,\mathrm{T})/|\mathrm{T}|) \quad (11)$$

where f (C_i , T)/ |T| is the probability that the selected case belongs to class C_i .

Land Use Land Cover (LULC) map

The supervised classification was used to prepare the LULC map for the change detection of the region. It is essential to estimate the area of various land cover classes such as Barren land, fallow, vegetation, settlements, water bodies. Such type of analysis is beneficial for the government and non-government agencies to implement policies and disaster mitigation strategies for the region. The following sections discuss the results and accuracy assessment of the classifier in detail.

IV. RESULTS AND DISCUSSION

In the year 2016, the study area was received 612mm rainfall in the monsoon season. Figure 3. shows the SPI value -0.46, which indicates the normal drought condition in the area. Therefore, the year was declared a typical drought year by the government of India. Consequently, 358mm rainfall was received in 2018 and 540mm in 2019, extremely low for two years. Such type of conditions is responsible for the water scarcity problems in the entire region. The SPI value -1.22 shows that the area was affected by severe drought conditions.

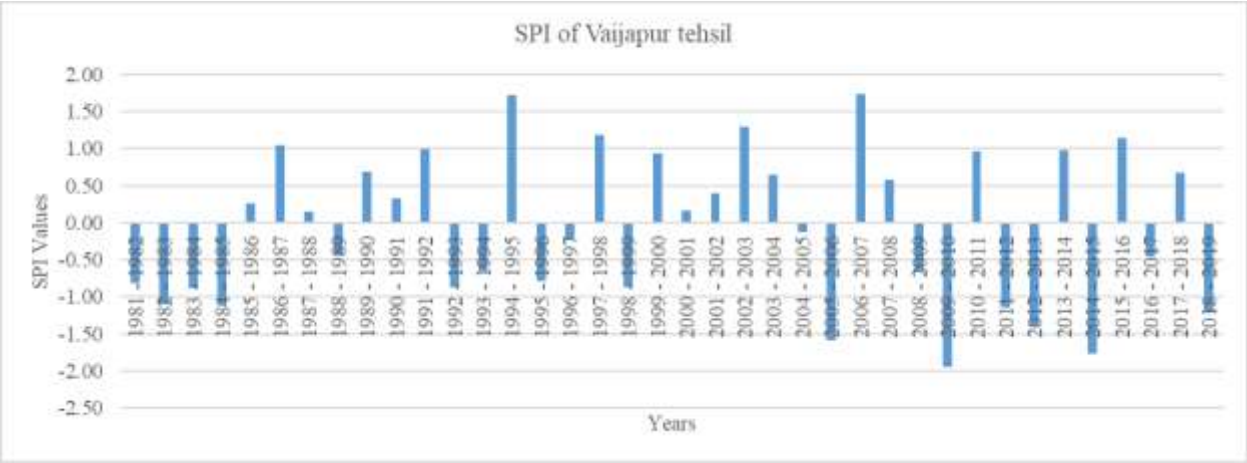

Figure. 3. SPI of year 1981-2019

According to the obtained result, minor changes are found in the settlement class and barren land class for 4 years. However, significant changes were found in the fallow land, vegetation and water class. These land cover classes were positively affected by water scarcity due to the deficient rainfall received in the region. In 2016, the fallow land was classified as 13465H, which 23771H highly increased in the year 2019. The sowing was partially completed in the region.

Furthermore, the sown crops were damaged due to the reduction in the soil moisture. Therefore, the agricultural productions were steepest declined in 2019. Total 55835H area was classified as the vegetation cover in 2016 and 31807H in 2019, which shows that the total 18599H vegetation cover was utterly damaged due to water stress. Similarly, the water bodies were classified as 11132H in 2016 due to sufficient rainfall in 2016. However, in the year 2019, the total 1840H area was classified by the classifier. Figure 5(A)(B) shows the classified map of the year 2016 and 2019, which clearly, shows the land cover changes in both the year.

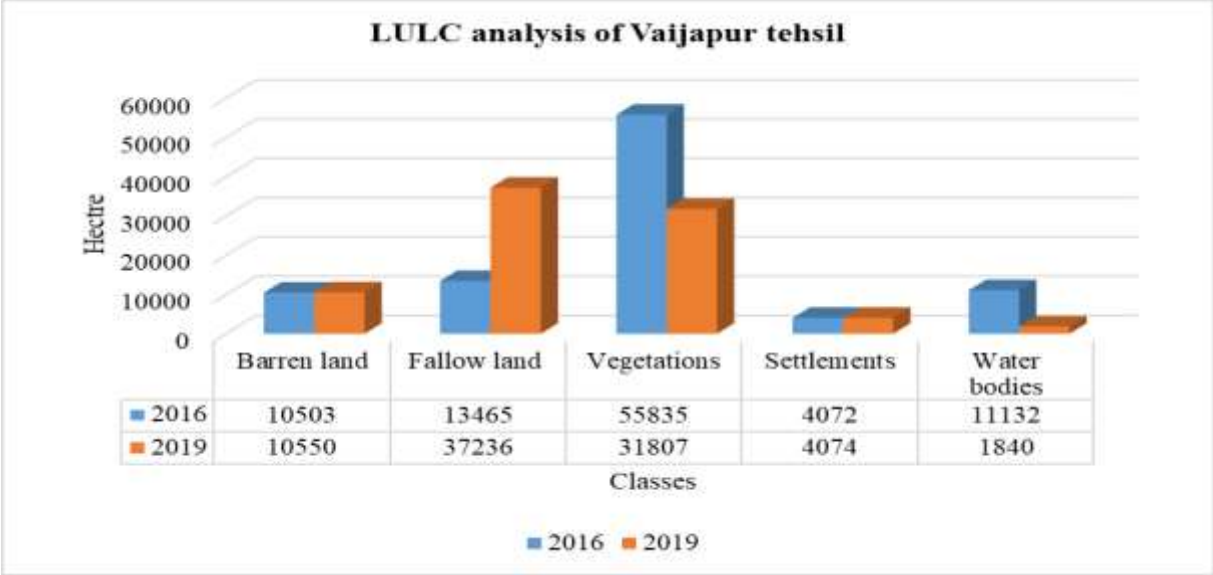

Figure. 4. LULC classification using Random Forest (RF) classifier

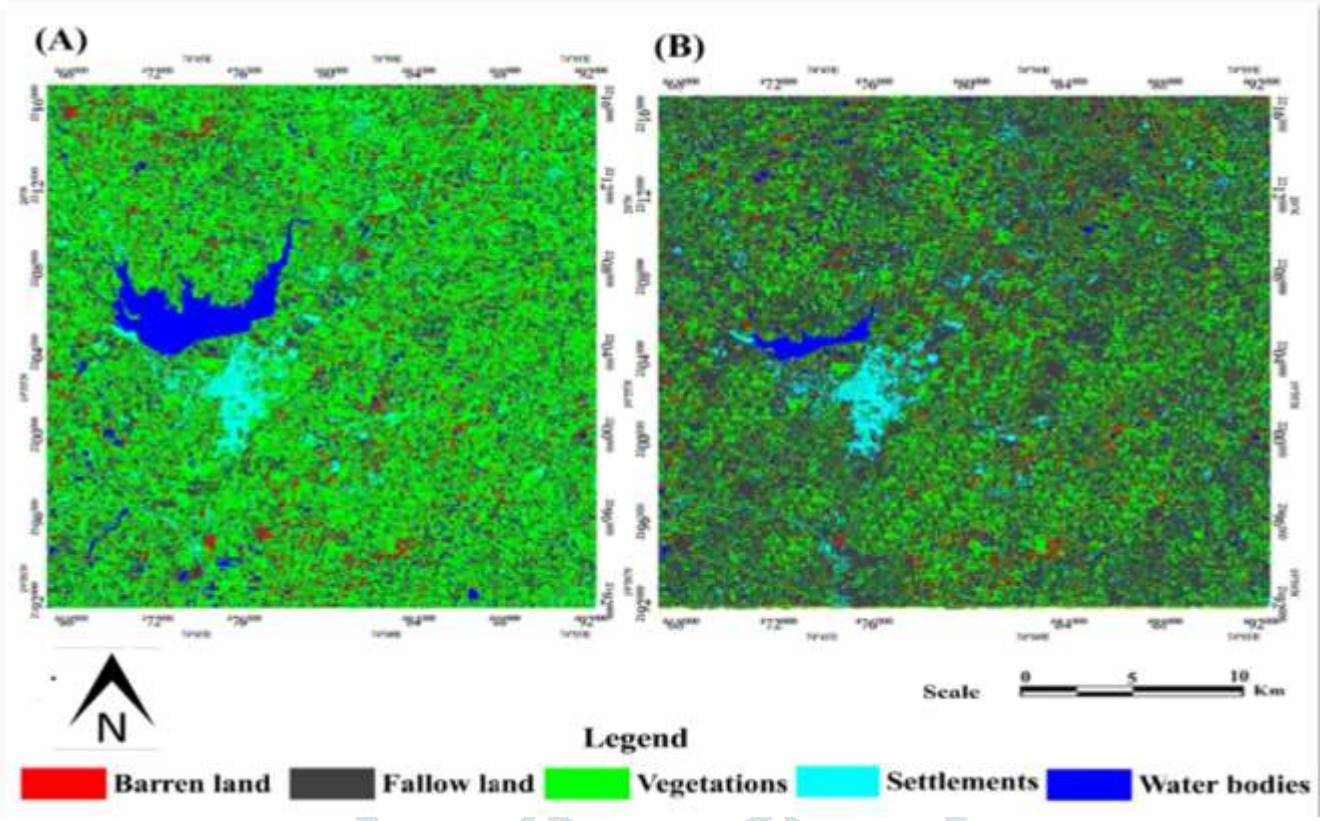

Figure. 5 (A). map of 2016 (B) LULC map of 2019

V. ACCURACY ASSESSMENT

The RF method's accuracy was accessed using a confusion matrix, which helps assess the classifier [32, 33, 34, 35]. In this research study, we have computed the Overall accuracy (OA), producer's accuracy (PA), user's accuracy (UA) and Kappa values of the year 2016-2019. It is shown in Table 2. The RF method's overall accuracy was determined to be approximately 92%, with a Kappa coefficient of 0.9 for 2016. Alternatively, corresponding values of accuracy and Kappa coefficient were 88.69% and 0.85, respectively, for 2019.

Table 2. Error matrix

Error matrix												
Class	Barren land		Fallow land		Vegetation		Settlements		Water bodies		Total	
	2016	2018	2016	2018	2016	2018	2016	2018	2016	2018	2019	2018
Barren land	130	120	16	26	0	0	0	0	0	0	146	146
Fallow land	15	35	150	130	0	0	0	0	0	0	165	165
Vegetation	10	5	0	0	140	145	0	0	0	0	150	150
Settlements	0	0	20	20	0	0	130	130	0	0	150	150
Water bodies	0	0	0	0	0	0	0	0	150	150	150	150
Total	155	160	186	176	140	145	130	130	150	150	761	761
PA (%)	83.87	75	80.64	73.86	100	100	100	100	100	100		
UA (%)	89.04	82.19	90.9	78.78	93.33	96.66	86.66	86.66	100	100		
Overall accuracy-2016 =91.98%, 2019= 88.69, Kappa Value-2016=0.9, 2019=0.85												

VI. CONCLUSION

In the present research study, the SPI and RF classification method. The SPI was calculated using the historical time series data of the 39 years from 1981-2019. The SPI value (1.14) for the year (2015-2016) shows normal drought. In contrast, the SPI value of -1.22 shows extreme drought because of the rainfall deficiency in the year 2019. The entire region was affected by the worst meteorological drought disaster in the year 2019 compared to 2016, which is mainly responsible for the area's water scarcity problems. The LULC analysis showed that the vegetation cover was affected due to deficient rainfall in the year 2019. The second dominant area was fallow land, which 23771H increased in 2019. These crops were damaged entirely due to scarcity of water for irrigation. Hence used RF method was successfully used for the land cover analysis of the study area.

ACKNOWLEDGEMENTS

The authors would like to thank the SARATHI fellowship for providing financial assistance to the MPhil degree. The author would like to thank Multispectral Laboratory, Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India.

REFERENCES

[1]. Bhuiyan, C., Singh, R. P., & Kogan, F. N. 2006. Monitoring drought dynamics in the Aravalli region (India) using different indices based on ground and remote sensing data. International Journal of Applied Earth Observation and Geoinformation, 8(4):289-302.

[2]. Hao, Z., & AghaKouchak, A. 2013. Multivariate standardized drought index: a parametric multi-index model. Advances in Water Resources. 57:12-18.

[3] AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., & Hain, C. R. 2015. Remote sensing of drought: Progress, challenges and opportunities. Reviews of Geophysics. 53(2):452-480.

[4]. Wilhelmi, O. V., & Wilhite, D. A. 2002. Assessing vulnerability to agricultural drought: a Nebraska case study. Natural Hazards. 25(1):37-58.

[5]. Vibhute, A.D., Dhumal, R., Nagne, A., Surase, R., Varpe, A., Gaikwad, S., Kale, K.V. and Mehrotra, S.C., 2017. Evaluation of soil conditions using spectral indices from hyperspectral datasets. In 2017 2nd International Conference on Man and Machine Interfacing (MAMI),IEEE, 1-6.

[6]. Hayes, M., Svoboda, M., Wall, N., & Widhalm, M. 2011. The Lincoln declaration on drought indices: universal meteorological drought index recommended. Bulletin of the American Meteorological Society. 92(4): 485-488.

[7]. Ghulam, A., Qin, Q., Teyip, T., & Li, Z. L. 2007. Modified perpendicular drought index (MPDI): a real-time drought monitoring method. ISPRS journal of photogrammetry and remote sensing. 62(2):150-164.

[8] Wong, G., Van Lanen, H. A. J., & Torfs, P. J. J. F. 2013. Probabilistic analysis of hydrological drought characteristics using meteorological drought. Hydrological Sciences Journal. 58(2):253-270.

[9] Van Loon, A. F., Van Lanen, H. A., Hisdal, H. E. G. E., Tallaksen, L. M., Fendeková, M., Oosterwijk, J., ... & Machlica, A. 2010. Understanding hydrological winter drought in Europe. Global Change: Facing Risks and Threats to Water Resources, IAHS Publ. 340:189-197.

[10] McKee, T. B. 1995. Drought monitoring with multiple time scales. In Proceedings of 9th Conference on Applied Climatology Boston.

[11] Naresh Kumar, M., Murthy, C. S., Sesha Sai, M. V. R., & Roy, P. S. 2009. On the use of Standardized Precipitation Index (SPI) for drought intensity assessment. Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling. 16(3):381-389.

[12] Livada, I., & Assimakopoulos, V. D. 2007. Spatial and temporal analysis of drought in Greece using the Standardized Precipitation Index (SPI). Theoretical and applied climatology. 89(3-4): 143-153.

[13] Hayes, M. J., Svoboda, M. D., Wiihite, D. A., Vanyarkho, O. V. 1999. Monitoring the 1996 drought using the standardized precipitation index. Bulletin of the American meteorological society. 80(3):429-438.

[14]. Rahmat, S. N., Jayasuriya, N., Adnan, M. S., & Bhuiyan, M. 2016. Analysis of spatio-temporal trends using standardised precipitation index (SPI). ARPN J Eng Appl Sci. 11(4): 2387-2392.

[15]. Yusof, F., Hui-Mean, F., Suhaila, J., Yusop, Z., & Ching-Yee, K. 2014. Rainfall characterisation by application of standardised precipitation index (SPI) in Peninsular Malaysia. Theoretical and applied climatology. 115(3-4):503-516

[16]. Rogan, J., & Chen, D. 2004. Remote sensing technology for mapping and monitoring land-cover and land-use change. Progress in planning. 61(4): 301-325.

[17]. Reis, S. 2008. Analyzing land use/land cover changes using remote sensing and GIS in Rize, North-East Turkey. Sensors. 8(10):6188-6202.

[18]. Vibhute, A.D., Kale, K.V., Gaikwad, S.V., Dhumal, R.K., Nagne, A.D., Varpe, A.B., Nalawade, D.B. and Mehrotra, S.C., 2020. Classification of complex environments using pixel level fusion of satellite data. Multimedia Tools and Applications, 1-33.

[19]. Gaikwad, S.V., Kale, K.V., Kulkarni, S.B., Varpe, A.B. and Pathare, G.N., 2015. Agricultural drought severity assessment using remotely sensed data: a review. International Journal of Advanced Remote Sensing and GIS, 4(1)1195-1203.

[20]. Chauhan, H. B., & Nayak, S. 2005. Land use/land cover changes near Hazira Region, Gujarat using remote sensing satellite data. Journal of the Indian society of Remote Sensing. 33(3):413-420.

[21]. Yan, F., Qin, Z., Li, M., & Li, W. 2006. Progress in soil moisture estimation from remote sensing data for agricultural drought monitoring. In Remote Sensing for Environmental Monitoring, GIS Applications, and Geology, International Society for Optics and Photonics. 6 (6366):636601.

[22]. Merem, E. C., & Twumasi, Y. A. 2008. Using geospatial information technology in natural resources management: the case of urban land management in West Africa. Sensors. 8(2): 607-619.

[23]. Bastiaanssen, W. G., Molden, D. J., & Makin, I. W. 2000. Remote sensing for irrigated agriculture: examples from research and possible applications. Agricultural water management. 46(2):137-155.

[24]. Ray, S. S., & Dadhwal, V. K. 2001. Estimation of crop evapotranspiration of irrigation command area using remote sensing and GIS. Agricultural water management. 49(3):239-249.

[25]. Lakhankar, T., Krakauer, N., & Khanbilvardi, R. 2009. Applications of microwave remote sensing of soil moisture for agricultural applications. International Journal of Terraspace Science and Engineering. 2(1):81-91.

[26]. Wang, L., & Qu, J. J. 2009. Satellite remote sensing applications for surface soil moisture monitoring: A review. Frontiers of Earth Science in China. 3(2):237-247.

[27]. Doraiswamy, P. C., Moulin, S., Cook, P. W., & Stern, A. 2003. Crop yield assessment from remote sensing. Photogrammetric engineering & remote sensing. 69(6):665-674.

[28]. Sobral, B. S., de Oliveira-Júnior, J. F., de Gois, G., Pereira-Júnior, E. R., de Bodas Terassi, P. M., Muniz-Júnior, J. G. R., ... & Zeri, M. 2019. Drought characterization for the state of Rio de Janeiro based on the annual SPI index: trends, statistical tests and its relation with ENSO. Atmospheric Research. 220:141-154.

[29]. Thom, H. C. 1958. A note on the gamma distribution. Monthly Weather Review 86(4):117-122.

[30]. Zhou, Han, Yuanbo Liu, and Yongwei Liu. "An approach to tracking meteorological drought migration." Water Resources Research 55, no. 4 (2019): 3266-3284.
[31]. Fawagreh, Khaled, Mohamed Medhat Gaber, and Eyad Elyan. "Random forests: from early developments to recent advancements." Systems Science & Control Engineering: An Open Access Journal 2, no. 1 (2014): 602-609.
[32]. Gaikwad, S.V., Vibhute, A.D., Kale, K.V., Dhumal, R.K., Nagne, A.D., Mehrotra, S.C., Varpe, A.B. and Surase, R.R., 2019. Drought Severity Identification and Classification of the Land Pattern Using Landsat 8 Data Based on Spectral Indices and Maximum Likelihood Algorithm. In Microelectronics, Electromagnetics and Telecommunications, Springer, Singapore, 517-524.
[33]. Gaikwad, S.V., Vibhute, A.D., Kale, K.V., Mehrotra, S.C., Dhumal, R.K., Varpe, A.B. and Surase, R.R., 2019. Identification and Classification of Water Stressed Crops Using Hyperspectral Data: A Case Study of Paithan Tehsil. In Proceedings of 2nd International Conference on Communication, Computing and Networking, Springer, Singapore, 911-919.
[34]. Nagne, A.D., Dhumal, R.K., Vibhute, A.D., Gaikwad, S., Kale, K. and Mehrotra, S., 2017, October. Land use land cover change detection by different supervised classifiers on LISS-III temporal datasets. In 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM), IEEE, 68-71.
[35]. Gaikwad, S.V. and Kale, K.V., 2015. Agricultural drought assessment of post monsoon season of Vaijapur Taluka Using Landsat8. Inter. J. of Res. in Eng. and Tech.(IJRET), 4(4).