International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

Extraction of Spread Surface Water Body using Supervised and Unsupervised Classification Techniques

B. Chandrababu Naik, B. Anuradha

***Abstract:** In this paper different classification techniques are applied to extract spread surface water area in the Nagarjuna sagar reservoir, Andhra Pradesh from Landsat-8 (OLI) image. In addition, the separability of reservoir features are tested to evaluate the thematic correctness of the classified data. This is to evaluate the application of a supervised and unsupervised classification techniques using the ERDAS software to extract the changes of surface water features for the period of 2014 to 2019. Furthermore, the statistical parameters are evaluated for the classification techniques. In supervised and unsupervised classification methods the minimum distance classifier gives better result (overall accuracy is 98.01%) than other classification methods. These obtained results are validated with ground truth data which is provided by Central Water-board Commission(CWC).*

***Keywords :** Landsat-8, surface water area, supervised and unsupervised classifications.*

I. INTRODUCTION

The continuous snapshots of the Earth's surface is provided by the Satellite-based remote sensing and Surface water dynamics [3] can also be done. The availability of spatial and temporal frequent observational data of physical attributes about the Earth's surface helps us to detect the dynamics of reservoirs at frequent time intervals. The recent studies on surface water bodies and flood areas shows that making use of satellite-based Earth Observation (EO) sensors to detect changes of surface water bodies are to be more focused.

The surface water area detection from multispectral satellite images shows that the water in infrared channels is less in reflectance than that of the water in other land cover types. [2] . In the past few decades [1] water indices, spectral bands, supervised and unsupervised classification techniques are widely used to detect the spread surface water bodies and it is also proven that it could separate water bodies from non-water bodies. By applying a threshold value, the water bodies are identified easily from non-water bodies in the multispectral satellite images.

In several time series analysis of wetlands, lakes, dams, reservoirs, rivers from the multispectral Landsat-8 (OLI) imagery is used to extract the water-bodies [4,5,6]. The surface water bodies in the reservoir of nagarjuna sagar are extracted using supervised and unsupervised classification techniques from Landsat 8 images for the period of 2014 to 2019. The Mean, Standard deviation, overall accuracy, and kappa coefficient of statistical parameters are evaluated.

II. STUDY AREA AND DATA

The reservoir of Nagarjuna sagar is located in Nalgonda District in Telangana and is constructed across the river of Krishna. The geographical area and its latitude 16°34'55.60"N to 16°56'44.95"N and longitude 78°24'13.97"E to 78°47'06.07"E as shown in fig. (1). The Landsat-8 data are downloaded from the USGS portal at free of cost [7,8,9], it contains the multispectral, spatial resolution is 30m, with a revisit period of 16 days, and provides 11 bands of information.

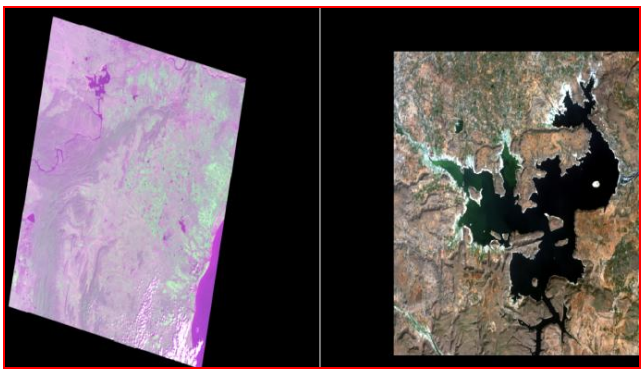

Fig.1 (a) Location of Nagarjuna Sagar reservoir (Landsat-8 (OLI) January, 2019) and (b) subset image

III. METHODOLOGY

The various classification techniques have developed to extract the surface water- bodies in multi spectral remote sensing imagery. The general techniques that are used as supervised and unsupervised classification techniques.

Manuscript received on February 10, 2020.
Revised Manuscript received on February 20, 2020.
Manuscript published on March 30, 2020.
* Correspondence Author
B. Chandrababu Naik*, Research Scholar, Department of Electronics and Communication Engineering, S. V University College of Engineering, Tirupati, Andhra Pradesh, India. Email: babunaikb@gmail.com
Dr B. Anuradha**, Professor, Department of Electronics and Communication Engineering, S. V University College of Engineering, Tirupati, Andhra Pradesh, India. Email: anubhuma@yahoo.com

In supervised classification, the most important techniques are Maximum likelihood classifier, Minimum distance classifier, and Spectral correlation mapper used to extract the water bodies. In unsupervised classification, the most familiar technique is K- means [11-14] algorithm is used to extract the water-bodies.

Unsupervised classification: The image classification is the procedure used is to classify image in pixels. For example, classes consist of water pixels, vegetation pixels, urban pixels. The unsupervised classification is a pure static analysis over multispectral imagery from without reference areas. Here the grouping of the pixels is following two procedures. 1. Beginning with individual pixels, 2. Beginning with all pixels. The flow chart of unsupervised classification techniques as shown in fig. (2).

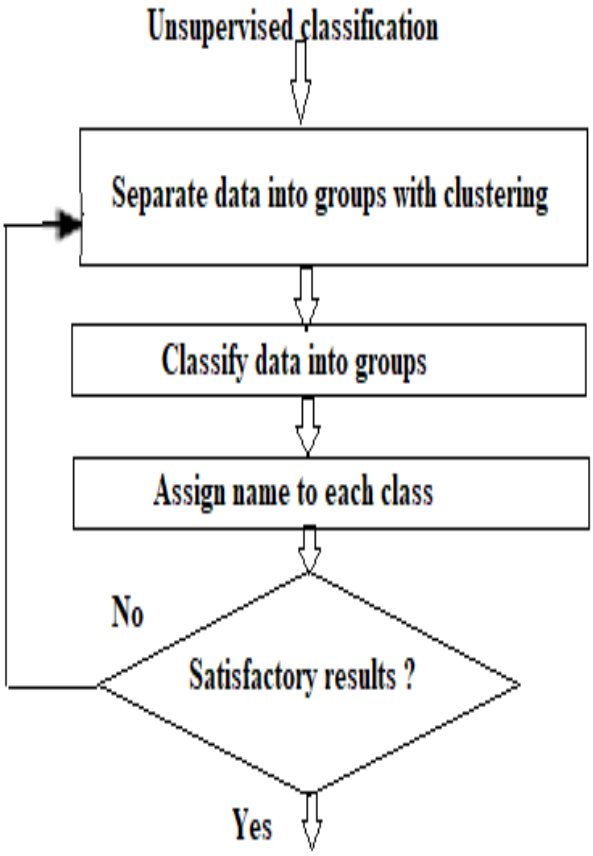

Fig.2 Flow chart for Unsupervised classification

K-Mean: The K- means algorithm [15] describes a particular number of disjoint clusters into the observation clusters, The algorithm is as follows;

1. The sample space is initially partitioned into K clusters, the observations are randomly allocated to the cluster.
2. Find the detachment among centroid of the cluster to observation cluster.
3. There is no observations are stimulated from one cluster to one more cluster to replicate the above two steps.

If clusters are becoming stable possible only when step 3 should be ended, and calculate the Euclidean distance (ED).

$$ED = \sqrt{\sum_{i=0}^{n}(xi - yi)^2} \quad .....(1)$$

Supervised Classification: It is a process of assigning the objects to their particular closest classes on the basis of mathematical dimensions derived from these objects. The flow chart for supervised classification as shown in fig. (3)

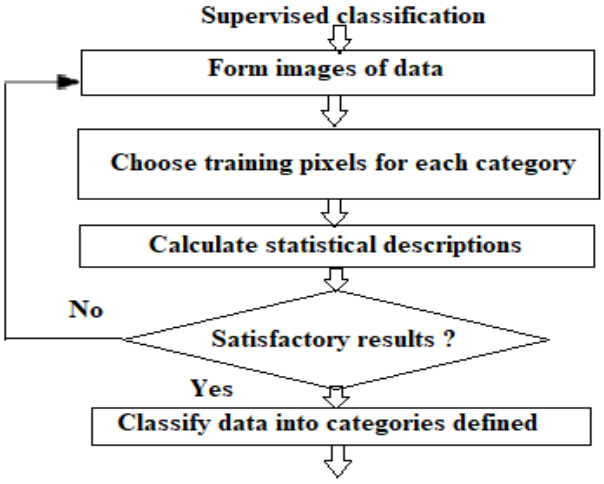

Fig.3 Flow chart for supervised classification

1. Maximum Likelihood Classification : In remote sensing applications the Maximum likelihood method is a familiar classifier used for supervised classification techniques for identifying delegate areas, nothing but training zones. It is the process to resolve the known class distribution for giving maximum statistic. Based on the Bayesian equation [10], this algorithm is computed for weighted distance (D) of unknown measurement vector (x) be in the right place to one of the known classes (Mc).

$$D = \ln(ac) - [0.5 \ln(|Covc|)] - [0.5 (X-Mc)T (Covc^{-1}) (X-Mc)] \quad .....(2)$$

Where, D = weighted distance, C = class, X = measurement vector, Mc = the mean vector, T= transposition function, ac= % probability that any candidate pixels, Covc= the covariance matrix of the pixels.

2. Minimum Distance Classification It describes the vector formed on the basis of known attributes and is represented by its mean vector [16,17]. This class utilizes the mean vectors for each class and Euclidean distance is estimated from each undefined pixel to mean vector for every class. This algorithm assigns the nearby values of the identical group of pixels, and it ignores the pixels when the distances exceeded a threshold value. The percentage of variance and covariance get increased and it depends on training pixels. The equation (3) discusses the information about the decision function for the closest mean.

$$d_i(x) = x^T \mu_i - 0.5\, \mu^T \mu_i \quad ..... (3)$$

Where x= pixels; $d_i(x)$ class with closest distance; μ_i= mean of the class. This classifier uses the classes whose correlation is zero and variance is same.

3. Spectral Correlation Mapper (SCM): SCM algorithm is the modified version of the Spectral Angle Mapper (SAM), where data is normalized and centered on the common between two spetral images[18]. The SCM model function provides several advantages. SAM is unable to discriminate the +ve and -ve correlations since it considers absolute values only. SAM algorithm is quantified only vector direction, not the magnitude of the effect of shading, but SCM algorithm solves these issues by normalizing each vector to the vector mean and reduces the shading effects [19].

The PCC (Pearsonian Correlation Coefficient) is the main difference of SAM and SCM operation to standardizes the data, the mean of x and Y is centralizing the vector.

The spectral correlation coefficient is defined with two n-dimensional vector x and y as

$$\mathrm{SCM} = \frac{\sum_i (x - x')\sum_i (y - y')}{\sqrt{\sum_i (x - x')^2}\sqrt{\sum_i (y - y')^2}} \quad \text{.....(4)}$$

Where x and y are expected values for the two vectors, x' and y' are sample mean respectively. SAM varies from 0 to 1 and SCM varies from -1 to 1. The spread surface water area extraction is done using supervised and unsupervised classification techniques from January-2014 to January-2019 as a time series analysis were performed.

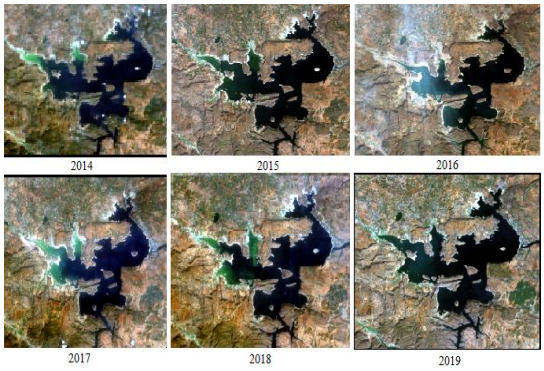

Fig.4 Landsat-8 input images from January, (2014, 2015, 2016, 2017, 2018 and 2019).

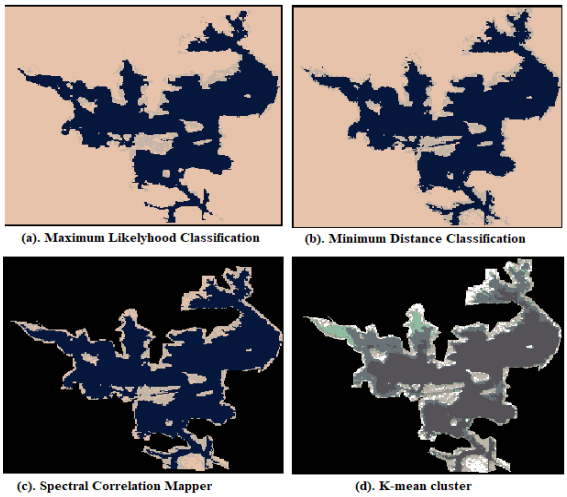

Fig.5 Output images of Supervised and Unsupervised Classification for 2014

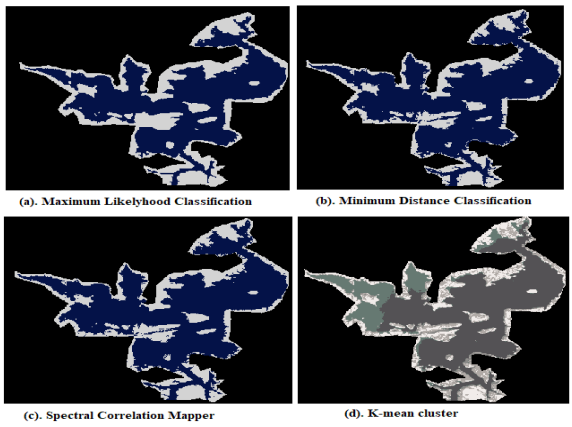

Fig.6 Output images of Supervised and Unsupervised Classification for 2015

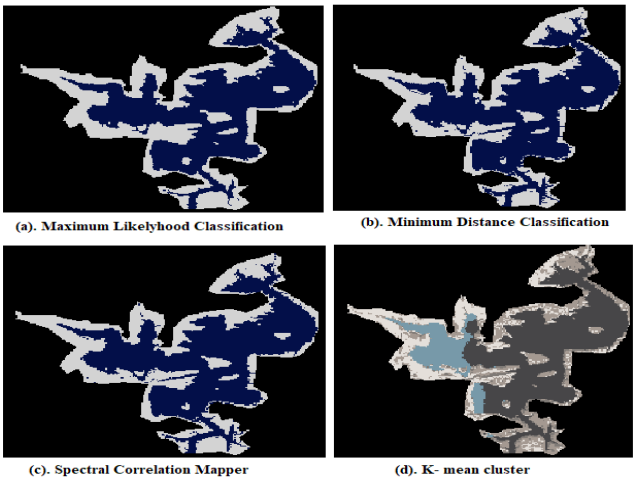

Fig.7 Output images of Supervised and Unsupervised Classification for 2016

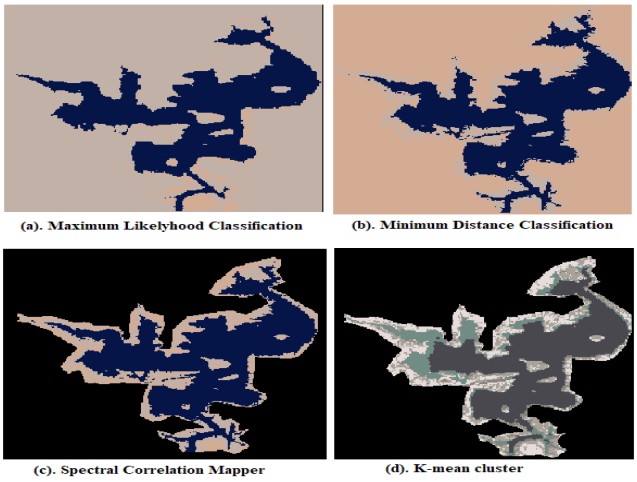

Fig.8 Output images of Supervised and Unsupervised Classification for 2017

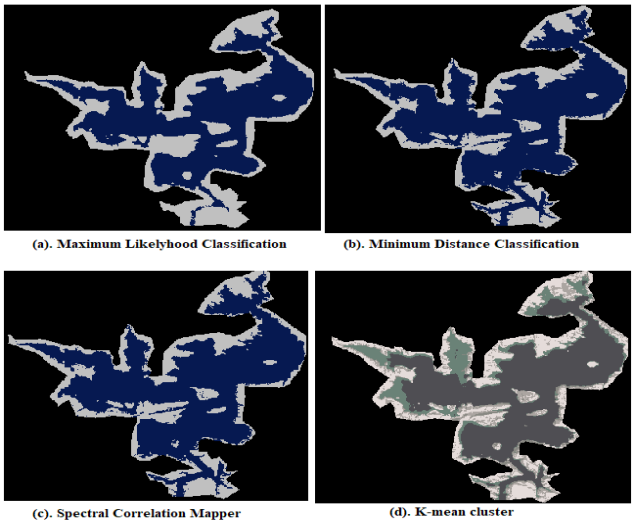

Fig.9 Output images of Supervised and Unsupervised Classification for 2018

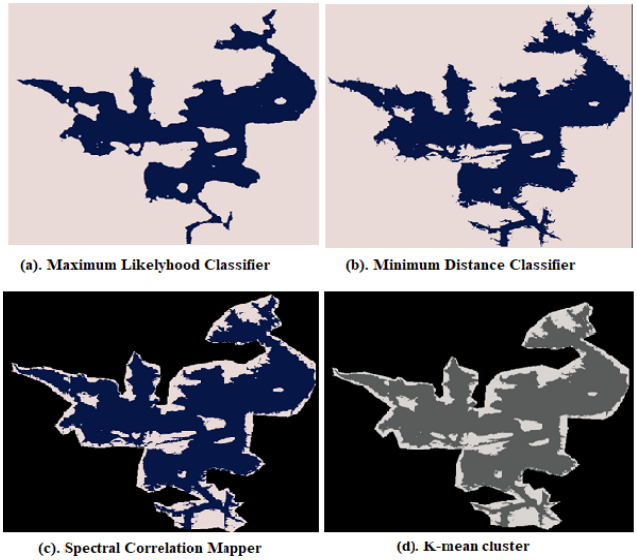

Fig.10 Output images of Supervised and Unsupervised Classification for 2019.

IV. RESULT AND DISCUSSION

The Landsat-8 images were used during the period of 2014 to 2019, subset the required part of the image act as an input image and pre- processing techniques are involved to reduce the spatial, spectral and Geo-corrections by using ERDAS software. Apply the input images for modeling of supervised and unsupervised classification techniques to extract the spread surface water area and its results as shown in fig. (4-10). The parametric calculation(mean, standard deviation, histogram pixels, and area in km^2) results for supervised and unsupervised classification techniques as shown in Table. (1). In supervised classification techniques obtained the results for surface water body extraction in the 2014 were Maximum likelihood classifier achieved 181.92 km^2, Minimum distance classifier achieved 194.92km^2, Spectral correlation mapper achieved 194.84 km^2. In unsupervised classification techniques K- mean cluster achieved 189.40 km^2. The Minimum distance classifier and Spectral correlation mapper provide the better results as compared to other classification techniques as shown in fig.(11).

Table. (1) The water-body extraction using Supervised and Unsupervised classification from 2014 to 2019

Years	Classification Techniques		Mean	Standard deviation	Histogram (pixels)	Area (km^2)
2014	Supervised Classification	1. Maximum Likelihood Classification	1.792	1.293	808553	181.92
		2. Minimum Distance Classification	1.874	1.313	866371	**194.92**
		3. Spectral Correlation Mapper	1.245	1.749	866069	194.84
	Unsupervised Classification	4. K-mean cluster	0.916	1.625	841841	189.40
2015	Supervised Classification	1. Maximum Likelihood Classification	1.405	1.790	727398	163.66
		2. Minimum Distance Classification	1.434	1.824	816439	**183.68**
		3. Spectral Correlation Mapper	1.434	1.823	815994	183.52
	Unsupervised Classification	4. K-mean cluster	0.890	1.496	774873	174.33
2016	Supervised Classification	1. Maximum Likelihood Classification	1.586	2.086	637165	135.17
		2. Minimum Distance Classification	1.626	2.131	697748	**156.97**
		3. Spectral Correlation Mapper	1.620	2.124	688803	154.95
	Unsupervised Classification	4. K-mean cluster	1.017	1.638	653270	146.98
2017	Supervised Classification	1. Maximum Likelihood Classification	2.591	1.242	625823	140.84
		2. Minimum Distance Classification	2.043	1.677	719777	**161.92**
		3. Spectral Correlation Mapper	2.032	1.668	717442	161.40
	Unsupervised Classification	4. K-mean cluster	0.974	1.590	687292	154.64
2018	Supervised Classification	1. Maximum Likelihood Classification	1.773	2.247	666860	150.04
		2. Minimum Distance Classification	1.813	2.295	789363	**177.59**
		3. Spectral Correlation Mapper	1.814	2.296	787987	177.28
	Supervised Classification	4. K-mean cluster	0.933	1.568	746063	167.85
2019	Supervised Classification	1. Maximum Likelihood Classification	1.619	2.123	680485	153.10
		2. Minimum Distance Classification	1.628	2.130	839013	**188.75**
		3. Spectral Correlation Mapper	1.664	2.108	837542	188.42
	Supervised Classification	4. K-mean cluster	0.503	0.692	800872	180.18

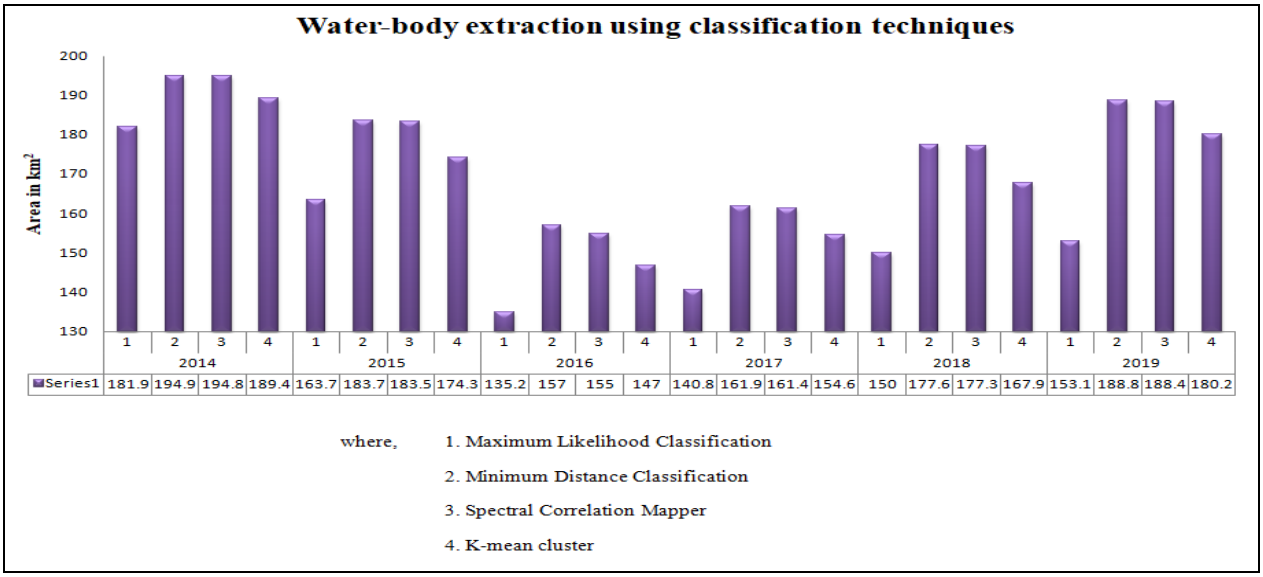

Fig.11 Graphical representation of water body extraction using classification techniques

The accuracy analyses have been done for supervised and unsupervised classifier during the period of 2014 to 2019. The overall accuracy and kappa coefficient were evaluated results as shown in fig. (12) and Table. (2). The Minimum distance classifier and Spectral correlation mapper achieved better results to extract the water body. The Minimum distance classifier overall accuracy is 98.01%, and a kappa value is 0.96. The Spectral correlation mapper overall accuracy is 97.84%, and a kappa value is 0.96. The Maximum likelihood classifier and K-mean cluster could not achieved better accuracy and kappa value as compared to the Minimum distance classifier and Spectral correlation mapper for extraction of spread surface water area in the reservoir nagarjuna sagar.

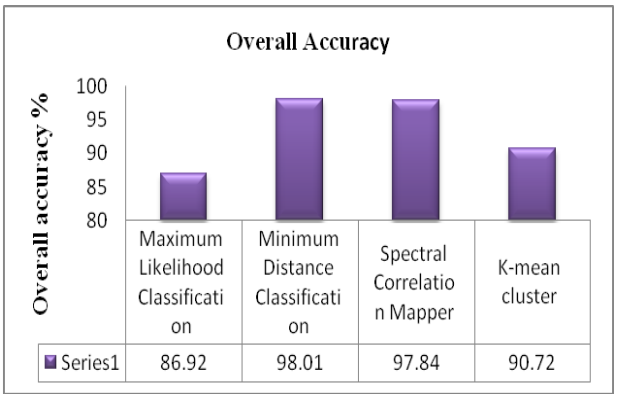

Fig.12 Graphical representation of accuracy assessment

Table. (2) A analyses of accuracy assessment

Classification techniques	Years (area in km^2)						Overall accuracy	Kappa coefficeent
	2014	2015	2016	2017	2018	2019		
Maximum Likelihood classifier	181.92	163.66	135.17	140.84	150.04	153.10	86.92	0.84
Minimum distance classifier	194.92	183.68	156.97	161.92	177.59	188.75	98.01	0.96
Spectral correlation mapper	194.84	183.52	154.95	161.40	177.28	188.42	97.84	0.96
k-mean cluster	189.40	174.33	146.98	154.64	167.85	180.18	90.72	0.89

Table.3 Comparison table of ground truth data with obtained results of water indices

Dates	Current reservoir levels (m)	Results of spread surface water area using classification techniques (km^2)			
		Maximum likelihood classifier	Minimum distance classifier	Spectral correlation mapper	K- mean cluster
14-01-2016	154.78	135.17	156.97	154.95	146.98
12-01-2017	157.58	140.84	161.92	161.40	154.64
11-01-2018	163.10	150.04	177.59	177.28	167.85
10-01-2019	165.02	153.10	188.75	188.42	180.18

.

The obtained results are validated with ground truth data, which is provided by Central Water-board Commission (CWC) real time data as shown in Table. (3). The CWC provides water level in terms of meters, but in this paper obtained results are extracted water area in terms of Sq. meters, it is concluded that the reservoir water level height is increased (in terms of meters) , then the water spread area is also increasing (in terms of Sq. meters) in the reservoir of nagarjuna sagar.

V. CONCLUSION

The dynamic changes of spread surface water area in the reservoir of nagarjuna sagar were evaluated using supervised and unsupervised classification techniques during the period of 2014 to 2019. The classification techniques (Maximum likelihood classifier, Minimum distance classifier, Spread correlation mapper, and K-mean cluster) were used for extraction of water area. The obtained results are in supervised classification techniques the Minimum distance classifier achieved an overall accuracy is 98.01%, and unsupervised classification k-mean cluster achieved an overall accuracy is 90.72%. The Minimum distance classifier provides better results as compared with other classifier techniques to extract the surface water body.

REFERENCES

1. Yang, Y., Liu, Y., Zhou, M., Zhang, S., Zhan, W., Sun, C., & Duan, Y. (2015). Landsat 8 OLI image based terrestrial water extraction from heterogeneous backgrounds using a reflectance homogenization approach. Remote Sensing of Environment, 171, 14–32. https://doi.org/ 10.1016/j.rse.2015.10.005
2. Huang, C., Chen, Y., Zhang, S., Wu, J (2018). Detecting, extracting, and monitoring surface water from space using optical sensors: A review. Rev. Geophys. 2018, 56, 333–360.
3. Vashistha, H.V., & Singh, P. (2015). Spatio-Temporal Landuse Change Monitoring of Surajpur Wetland through Remote Sensing and GIS. 16th ESRI India User Conference. New Delhi.
4. Haque, M.I., & Basak, R. (2017). Land cover change detection using GIS and remote sensingtechniques: A spatio-temporal study on Tanguar Haor, Sunamganj, Bangladesh. The Egyptian Journal of Remote Sensing and Space Science, 20, 251-263
5. Emch, M., & Peterson, M. (2006). Mangrove forest cover change in the Bangladesh Sundarbans from 1989-2000: A remote sensing approach. Geocarto International, 21, 5-12
6. Elsahabi, M., Negm, A., & El Tahan, A.H.M. (2016). Performances evaluation of surface water areas extraction techniques using Landsat ETM+ data: case study Aswan High Dam Lake (AHDL). Procedia Technology, 22, 1205-1212
7. U.S. Geological Survey. USGS Global Visualization Viewer. Available online: http://glovis.usgs.gov (accessed on 19 March 2014)
8. Goodwin, N.R.; Collett, L.J.; Denham, R.J.; Flood, N.; Tindall, D. Cloud and cloud shadow screening across Queensland, Australia: An automated method for Landsat TM/ETM+ time series. Remote Sens. Environ. 2013, 134, 50–65.
9. Verpoorter, C.; Kutser, T.; Tranvik, L. Automated mapping of water bodies using Landsat multispectral data. Limnol. Oceanogr. 2012, 10, 1037–1050
10. Otukei, J.; Blaschke, T. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. Int. J. Appl. Earth Obs. Geoinf. 2010, 12, S27–S31.
11. Rundquist, D.C.; Lawson, M.P.; Queen, L.P.; Cerveny, R.S. The relationship between summer-season rainfall events and lake-surface area. J. Am. Water Resour. Assoc. 1987, 23, 493–508.
12. Lu, D.; Weng, Q. A survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. 2007, 28, 823–870.
13. Otukei, J.; Blaschke, T. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. Int. J. Appl. Earth Obs. Geoinf. 2010, 12, S27–S31
14. Frazier, P.S.; Page, K.J. Water body detection and delineation with Landsat TM data. Photogramm. Eng. Remote Sens. 2000, 66, 1461–1468.
15. G. Hamerly and C. Elken. Altenatives to the K-mean algorithm that find better clustering. In Proc. of the 11th Intl, Conf. on Information and Knowledge Management, Pages 600-607, Mclean, Viginia, 2002
16. A.G.Wacker, D.A.Landgrebe-"Minimum Distance Classification in Remote Sensing"- The Laboratory for Applications of Remote Sensing Purdue University Lafayette, Indiana.
17. James R. Carr (Editor), 1999. Classification of Digital Image Texture using Variograms. Advances in Remote sensing and GIS Analysis. John Wiley & Sons Ltd.
18. Conel, James E; Hoover, G; Nolin, A; Alley, R; Margolis, J., 1992. Empirical Relationships Among Semi- Arid Landscape Endmembers Using the Spectral Angle Mapper (SAM) Algorithm, Summaries of the 4nd Annual JPL Airborne Geoscience Workshop, JPL Pub-92-14, AVIRIS Workshop. Jet Propulsion Laboratory, Pasadena, CA, pp. 150-151
19. O. A. de Carvalho Jr and P. R. Meneses, "Spectral Correlation Mapper (SCM): An Improvement on the Spectral Angle Mapper (SAM)," Summ. 9th JPL Airborne Earth Sci. Work. JPL Publ., no. equation 2, p. 9 pages, 2000.

AUTHORS PROFILE

B. Chandrababu Naik, is a Research Scholar at Department of Electronics and Communication Engineering , Sri Venkateswara University, Tirupati. He is B. Tech in Electronics and Communication Engineering from S V University, TIRUPATI, and M.Tech in Communicatio System from S V University, Tirupati. He has Teaching experience of 4 years. His Areas of Interest is Digital Image Processing for Satellite Imagery.

Dr B. Anuradha, is currently working as a professor at the Department of Electronics and Communication Engineering, Sri Venkateswara University College of Engineering,Tirupati,Andhra Pradesh.She has a teaching experience of 25 Years and has many technical publications in national/international journals. Her areas of interest include Digital Image Processing and biomedical signal processing..