

Original Research Article

<https://doi.org/10.20546/ijcmas.2018.707.025>

Accuracy Assessment of Supervised and Unsupervised Classification using Landsat-8 Imagery of D-7 Shahapur Branch Canal of UKP Command Area Karnataka, India

Ayyanna^{1*}, B.S. Polisgowdar², M.S. Ayyanagowdar², Anilkumar T. Dandekar³, G.S. Yadahalli⁴ and M.A. Bellakki⁵

¹Agri. Engg, Department of SWE UAS, Raichur, India

²Department of IDE, ³Department of SWE, CAE, UAS, Raichur, India

⁴Department of Agronomy, AC, Raichur, India

⁵Department of Soil Science and Agricultural Chemistry, AC, Kalaburgi, India
Department of Soil and Water Conservation Engineering, College of Agricultural Engineering, University of Agricultural Sciences, Raichur - 584 104, India

*Corresponding author

ABSTRACT

Keywords

Landsat-8,
Unsupervised and
Supervised
classification,
Command area

Article Info

Accepted:

04 June 2018

Available Online:

10 July 2018

An investigation was conducted from November 2015 to May 2017 at Shahapur and Shorapur taluk, Yadgir district, to accuracy assessment of supervised and unsupervised classification for land use land cover using Landsat-8 imagery of D-7 Shahapur branch canal of UKP command area, Karnataka. Accuracy assessment of supervised land classification for post and pre-monsoon was carried out for the study period and the values are presented in Table 5 to 8. The technique used for this was called error of confusion matrix. The overall accuracy for both post and pre-monsoon ranged from 86.66 per cent to 92 per cent and Kappa co-efficient ranged from 0.81 to 0.89. The range value of classification was due to intermixing of some classes Megersa (2012), that any accuracy assessment value more than 75 per cent will be considered acceptable Giles (2008). Similarly for kappa co-efficient more than 0.80 indicate moderate and less than 0.40 indicate poor classification. Kappa co-efficient for this study were within the range expressed as good classification, hence classification yielded good results and comparable to other studies.

Introduction

When identifying land use or land cover for a given area of interest, two common approaches to classify each pixel in an image are supervised classification and unsupervised classification. In supervised classification, an analyst uses previously acquired knowledge of

an area, or a prior knowledge, to locate specific areas, or training sites, which represent homogeneous samples of known land use land cover types. Based on statistics of these training sites, each pixel in an image is then assigned to a user-defined land use land cover type (crop land, grassland, salt affected and waterlogged *etc.*). Unsupervised

classification is useful for scenes in which land cover is not well-known or undefined. Computer algorithms group similar pixels into various spectral classes which the analyst must then identify and combine into information classes (Jensen 2005, Thomson *et al.* 1998). Both approaches of classification have strengths and weaknesses associated with the physical execution of the classification process and with the final result of the analysis. Instead, the intent of this research was to determine situations where one classification method is predisposed to be more accurate than the other based on contributing factors such as training site selection, spectral reflectance properties, and land cover composition. Because remote sensing is a powerful tool for studying geospatial phenomena, land use and land cover studies make frequent use of remotely sensed imagery. There are a wide variety of applications for land use and land cover studies to include natural disaster mapping (Borghuis *et al.*, 2007). Possessing accurately classified imagery is paramount to these studies and can affect decisions regarding land development.

If there are measurable and predictable imagery characteristics that would increase the likelihood of one classification method possessing higher accuracy than others, researchers could save a great amount of time and manpower by utilizing this knowledge before conducting any analysis.

There is extensive literature that supports the individual use of both supervised and unsupervised classification approaches in land use and land cover studies. Additionally, there are many studies where both classification approaches are compared to determine which approach is more accurate (Borghuis *et al.*, 2007), while a greater number of studies found the converse to be true (Alrababah and Alhamad, 2006; Bahadur, 2009; Mukherjee and Mukherjee, 2009; Trisurat *et al.*, 2000).

Materials and Methods

The study area is located between 16° 29' N to 16° 39' N latitude and 76°43' E to 76°54' E longitude. It is located 45 km away from the district headquarters and location map as showed below Figure 1.

The gross command area of D-7 distributary is 15647.49 ha with a cultivable command area of 10322 ha and with a discharge capacity of 5.876 cumecs. The canal command covers parts of three toposheets of survey of India namely E43W10, E43W14 and E43W15 in the scale of 1:50000.

The climate of the region is generally dry. The larger variations in rainfall from year to year both in quantity and distribution throughout the season render the region to drought. As per Koppen's classification, the regions climate is semi-arid. The command area experiences an average annual rainfall of 656 mm. Though the total rainfall is not high, the area benefits both from the South-West and North-East monsoons. The region is having dry climate. The months from December to May are the driest.

Results and Discussion

Classification of image for land use and land cover

The approach is based upon the classification of the remote sensing data for classification of land use and land cover in the study area. Two different type of classification *i.e.* unsupervised and supervised classification were adopted.

Processing of remote sensing data

The data of landsat-8 for four images were used for the present study. Initially a false colour composite (FCC) of bands 5, 4 and 3 was prepared and visualised. Then, each individual band was visualised one by one while using

the temporal satellite data of the same area and the images were registered to the base map. The map imported from ArcGIS was considered as the base map. For carrying out registration, ground control points (GCPs) like crossing of rivers, canals, sharp turns in the rivers, bridges, road/canal intersection were located on the images. The operational flow process for the work is shown in Figure 2.

Unsupervised classification

Unsupervised classification is a method which examines a large number of unknown pixels and divides into a number of classes based on natural groupings present in the image value. Unsupervised classification does not require analyst-specified training data. The basic premise is that within a given cover type should be close together in the measurement space, where data in different classes should be comparatively well separated. The classes that results from unsupervised classification are spectral classes which are based on natural grouping of the image values, the identity of spectral class will not be initially known, must compare classified data from reference data (such as larger scale imagery, maps, or site visit) to determine the identity and information values of the spectral classes.

Supervised classification

The supervised classification method is based upon three band of landsat-8 Band 3 (Green), Band 4 (Red) and Band 5 (NIR) with FCC as the background map. Supervised training is closely controlled by the analyst. In this process, the pixels that represent patterns or land cover features that are recognizable were selected, or that can be identified with the help of ground truth data. Knowledge of the data, and of the classes desired, is therefore required before classification. By identifying these patterns, the computer system is instructed to identify pixels with similar characteristics. The pixels identified by the training samples

are analyzed statistically to form what are referred to as signatures. After the signatures are defined, the pixels of the image are sorted into classes based on the signatures by use of a classification decision rule.

The decision rule is a mathematical algorithm that, using data contained in the signature, performs the actual sorting of pixels into distinct class values. If the classification is accurate, the resulting classes represent the categories within the data that are originally identified with the training samples. Supervised classification can be used as a term to refer to a wide variety of feature extraction

Accuracy assessment

The accuracy assessments of both techniques were made through a confusion or error matrix. A confusion matrix contains information about actual and predicted classifications done by a classification system. The pixel that has been categorized from the image was compared to the same site in the field. The result of an accuracy assessment typically provides the users with an overall accuracy of the map and the accuracy for each class in the map. The percentage of overall accuracy was calculated using following formula.

Land use and land cover classification

The land use/land cover (LU/LC) classification was carried out using unsupervised and supervised classification. Interactive Self-Organizing Data Analysis (ISODATA) Technique for unsupervised classification and Maximum Likelihood Classifier (MLC) algorithm for supervised classification was used in the present analysis.

Unsupervised classification

For LU/LC classification, the ERDAS Imagine classifier module has been used. In

ERDAS, unsupervised classification, the user input the number of clusters desired and a confidence threshold (usually 0.95).

The computer will then build clusters iteratively, meaning that with each new iteration, the clusters become more and more refined. The iteration stops when the confidences level is reached. After the clusters are built, the analyst must select the land cover classes and then assign each cluster to appropriate class. In this study, unsupervised classification with 30 clusters for images of November-2015, May-2016, November-2016 and May-2017 and 95 per cent confidences threshold has been performed.

The result was image with groups of pixels each represented by a different colour. Now we have to highlight each cluster one at a time and then determine which of the class belongs to, by interpreting the original multispectral image. Then each cluster colour has to be changed to appropriate one. Finally the image was recorded into the number of class as it is presented in Figure 3. The abstract of the area occupied by different classes for the pre and post monsoon are presented in Table 1 and 2 respectively.

From Table 1 it is observed that the comparison of unsupervised classification for November 2015 and 2016 total number of classes were eight. Among these classes cropped area was highest having total area of 2551.77 and 3700.53 ha and minimum area was water body 649.98 and 713.07 ha which accounts for 16.31, 23.65 and 4.15, 4.56 per cent of the cultivable command area respectively during November 2015 and 2016. It was also observed that salt affected area was 1567.44 and 1232.91 ha and waterlogged area was 2033.28 and 1703.07 ha which accounts for 10.02, 7.88 and 12.99, 10.88 per cent of the cultivable command area during November 2015 and 2016 respectively.

Table 2 presents the comparison of unsupervised classification for May 2016 and 2017 in which it was observed that total number of classes were eight and among these classes fallow land area was highest having total area of 2388.87 (May 2016) and 4706.19 ha (May 2017) and minimum area was settlement (1336.86 and vegetation 569.34 ha) which accounts for 15.27, 30.08 and 8.54, 3.64 per cent of the cultivable command area respectively during May 2016 and 2017. Also, salt affected area was 1558.53 and 1595.52 ha and waterlogged area was 2026.89 and 1149.66 ha which accounts for 9.96, 10.20 and 12.95, 7.35 per cent of the cultivable command area respectively during May 2016 and 2017.

Supervised classification

In supervised classification, spectral signatures files were created called as training sites based on visual interpretation with the prior knowledge of image interpretation techniques and ground truthing.

Once the training sites are developed, computer can use this information, along with the various images of different band widths, to create spectral signatures from the specified areas. The abstract the areas occupied by different classes for the said period are presented in Table 3 and 4 respectively.

From Table 3 comparison of supervised classification for Nov-2015 and Nov-2016 indicated that, the total number of classes were eight and among these classes cropped area was highest having total area of 6460.62 and 8261.55 ha which accounts 41.28 and 52.79 per cent and minimum area was water body 146.79 and 131.49 ha which accounts for 0.93 and 0.84 per cent of the cultivable command area respectively.

During the same period salt affected area was 1797.03 and 1762.11ha which accounts 11.48 and 11.26 per cent and waterlogged area was found 1441.98 and 1489.77 ha which accounts for 9.21 and 9.52 per cent of the cultivable command area respectively.

Similarly during May 2016 and 2017 it was observed that (Table 4) total number of classes were eight and among these classes fallow land area was highest having total area of 4982.67 and 5815.26 ha which accounts 31.84 and 37.16 per cent and minimum area was waterlogged 52.87 and 53.46 ha which accounts for 0.33 and 0.34 per cent of the cultivable command area respectively. Also the salt affected area was 3486.96 and 3602.86 ha and waterlogged area was found 52.87 and 53.46 ha which accounts for 22.28 and 23.02 per cent of the cultivable command area during May 2016 and 2017 respectively.

Accuracy assessment of satellite image classification

The confusion matrix showing the overall

accuracy along with the Kappa co-efficient for supervised classification for post-monsoon 2015 and 2016 and pre-monsoon 2016 and 2017 are presented in Table 5 to 8. From Table 5 during November 2015, indicates accuracy of land use/land cover classification shows that out of eight class water body, waterlogged and rock out crop have user accuracy 100 per cent. The producer accuracy for water body and settlements are 100 per cent. The least user accuracy was found for vegetation 77.77 per cent followed by settlements 83.33 per cent and salt affected area 83.33 per cent. Kappa co-efficient was found 0.86.

From confuse matrix of November, 2016 (Table 6) it is observed that accuracy of land use/land cover classification out of eight class's water body area, waterlogged area and settlements and rock-out crop have user accuracy of 100 per cent. The least user accuracy was found for salt affected area 80 per cent followed by fallow land 83.33 per cent and Kappa co-efficient was found 0.89.

Table.1 Comparison of unsupervised classification for November-2015 and 2016

Sl. No	Land cover	November 2015		November 2016		Change of percentage (%)
		Area (ha)	Percentage of area (%)	Area (ha)	Percentage of area (%)	
1	Crop land	2551.77	16.31	3700.53	23.65	45.02
2	Fallow land	2531.25	16.18	1524.24	9.74	-39.78
3	Rockout crop	1897.74	12.13	2065.95	13.20	8.86
4	Salt affected area	1567.44	10.02	1232.91	7.88	-21.34
5	Settlements	2082.24	13.31	2196.18	14.04	5.47
6	Vegetation	2333.79	14.91	2511.54	16.05	7.62
7	Water body	649.98	4.15	713.07	4.56	9.71
8	Waterlogged	2033.28	12.99	1703.07	10.88	-16.24
	Total	15647.49	100	15647.49	100	

Table.2 Comparison of unsupervised classification for May-2016 and May-2017

Sl. No	Land cover	May 2016		May 2017		Change percentage (%)
		Area (ha)	Percentage of area (%)	Area (ha)	Percentage of area (%)	
1	Crop land	2369.07	15.14	1795.05	11.47	-24.23
2	Fallow land	2388.87	15.27	4706.19	30.08	97.00
3	Rockout crop	2325.69	14.86	3075.66	19.66	32.25
4	Salt affected area	1558.53	9.96	1595.52	10.20	2.37
5	Settlements	1336.86	8.54	1990.08	12.72	48.86
6	Vegetation	2274.12	14.53	569.34	3.64	-74.96
7	Water body	1367.46	8.74	765.99	4.90	-43.98
8	Waterlogged	2026.89	12.95	1149.66	7.35	-43.28
	Total	15647.49	100	15647.49	100	

Table.3 Comparison of supervised classification for November-2015 and 2016

Sl. No.	LU/LC	November 2015		November 2016		Change percentage (%)
		Area (ha)	Percentage of area (%)	Area (ha)	Percentage of area (%)	
1	Crop land	6460.62	41.28	8261.55	52.79	27.88
2	Fallow land	2151.54	13.75	153.54	0.98	-92.86
3	Rockout crop	1834.11	11.72	2125.89	13.58	15.91
4	Salt affected area	1797.03	11.48	1762.11	11.26	-1.94
5	Settlements	498.81	3.187	510.66	3.26	2.38
6	Vegetation	1316.61	8.41	1212.48	7.74	-7.91
7	Water body	146.79	0.93	131.49	0.84	-10.42
8	Waterlogged	1441.98	9.21	1489.77	9.52	3.31
	Total	15647.49	100	15647.49	100	

Table.4 Comparison of supervised classification for May-2016 and May-2017

Sl. No	LU/LC	May 2016		May 2017		Change percentage (%)
		Area (ha)	Percentage of area (%)	Area (ha)	Percentage of area (%)	
1	Crop land	1936.72	12.37	1206.82	7.71	-37.69
2	Fallow land	4982.67	31.84	5815.26	37.16	16.71
3	Rockout crop	3119.13	19.93	3108	19.86	-0.36
4	Salt affected area	3486.96	22.28	3602.86	23.02	3.32
5	Settlements	530	3.38	538	3.438	1.51
6	Vegetation	1444.14	9.22	1230.39	7.86	-14.80
7	Water body	95	0.60	92.7	0.59	-2.42
8	Waterlogged	52.87	0.33	53.46	0.34	1.12
	Total	15647.49	100	15647.49	100	

Table.5 Confuse matrix of supervised classification of November-2015

Land cover	Water body	Waterlogged area	Crop land	Fallow land	Vegetation	Settlements	Salt affected area	Rockout crop	Total	User accuracy
Water body	1	0	0	0	0	0	0	0	1	100
Waterlogged area	0	3	0	0	0	0	0	0	3	100
Crop land	0	1	39	1	0	0	0	0	41	95.12
Fallow land	0	0	1	11	0	0	0	1	13	84.61
Vegetation	0	0	0	0	7	0	2	0	9	77.77
Settlements	0	0	1	0	0	5	0	0	6	83.33
Salt affected area	0	0	0	0	1	0	5	0	6	83.33
Rockout crop	0	0	0	0	0	0	0	3	3	100
Total	1	4	41	12	8	5	7	4	82	
Producer accuracy	100	75	95.12	91.67	87.5	100	71.42	75	OA	90.00
Kappa co-efficient = 0.86										

Table.6 Confuse matrix of supervised classification of November-2016

Land cover	Water body	Waterlogged area	Crop land	Fallow land	Vegetation	Settlements	Salt affected area	Rockout crop	Total	User accuracy
Water body	1	0	0	0	0	0	0	0	1	100
Waterlogged area	0	3	0	0	0	0	0	0	3	100
Crop land	0	1	23	0	0	0	1	0	25	92
Fallow land	0	0	0	10	0	1	0	0	11	90.90
Vegetation	0	0	0	1	11	0	0	0	12	91.67
Settlements	0	0	0	0	0	5	0	0	5	100
Salt affected area	0	0	0	0	1	0	4	0	5	80
Rockout crop	0	0	1	0	0	0	0	12	13	92.30
Total	1	4	24	11	12	6	5	12	75	
Producer accuracy	100	75	95.83	90.90	91.67	83.33	80	100	OA	92.00
Kappa co-efficient = 0.89										

Table.7 Confuse matrix of supervised classification of May-2016

Land cover	Water body	Waterlogged area	Crop land	Fallow land	Vegetation	Settlements	Salt affected area	Rockout crop	Total	User accuracy
Water body	1	0	0	0	0	0	0	0	1	100
Waterlogged area	0	4	1	0	0	0	0	0	5	80
Crop land	0	1	30	0	0	0	1	0	32	93.75
Fallow land	0	0	1	7	1	0	0	0	9	77.77
Vegetation	0	0	0	0	3	0	0	0	3	100
Settlements	0	0	0	0	0	8	0	1	9	88.88
Salt affected area	0	0	1	1	0	0	5	0	7	71.42
Rockout crop	0	0	1	0	0	1	0	7	8	77.77
Total	1	5	34	8	4	9	6	7	75	
Producer accuracy	100	80	88.23	87.5	75	88.88	83.33	87.5	OA	86.66
Kappa co-efficient = 0.81										

Table.8 Confuse matrix of supervised classification of May-2017

Land cover	Water body	Waterlogged area	Crop land	Fallow land	Vegetation	Settlements	Salt affected area	Rock out crop	Total	User accuracy
Water body	1	0	0	0	0	0	0	0	1	100
Waterlogged area	0	6	0	0	0	0	0	0	6	100
Crop land	0	0	29	1	0	0	0	0	30	96.67
Fallow land	0	0	0	7	0	1	0	1	9	77.78
Vegetation	0	0	1	0	5	0	1	0	7	71.42
Settlements	0	0	0	0	0	5	0	0	5	100
Salt affected area	0	0	0	0	2	0	9	1	12	75
Rock out crop	0	0	0	0	0	1	0	10	11	90.90
Total	1	6	30	8	7	7	10	12	81	
Producer accuracy	100	100	96.67	87.5	71.42	71.42	90	83.33	OA	88.89
Kappa co-efficient = 0.84										

Fig.2 Operational flow chart

$$\text{Overall accuracy} = \frac{\text{Total number of correct samples}}{\text{Total number of samples}} \times 100 \quad \dots (1)$$

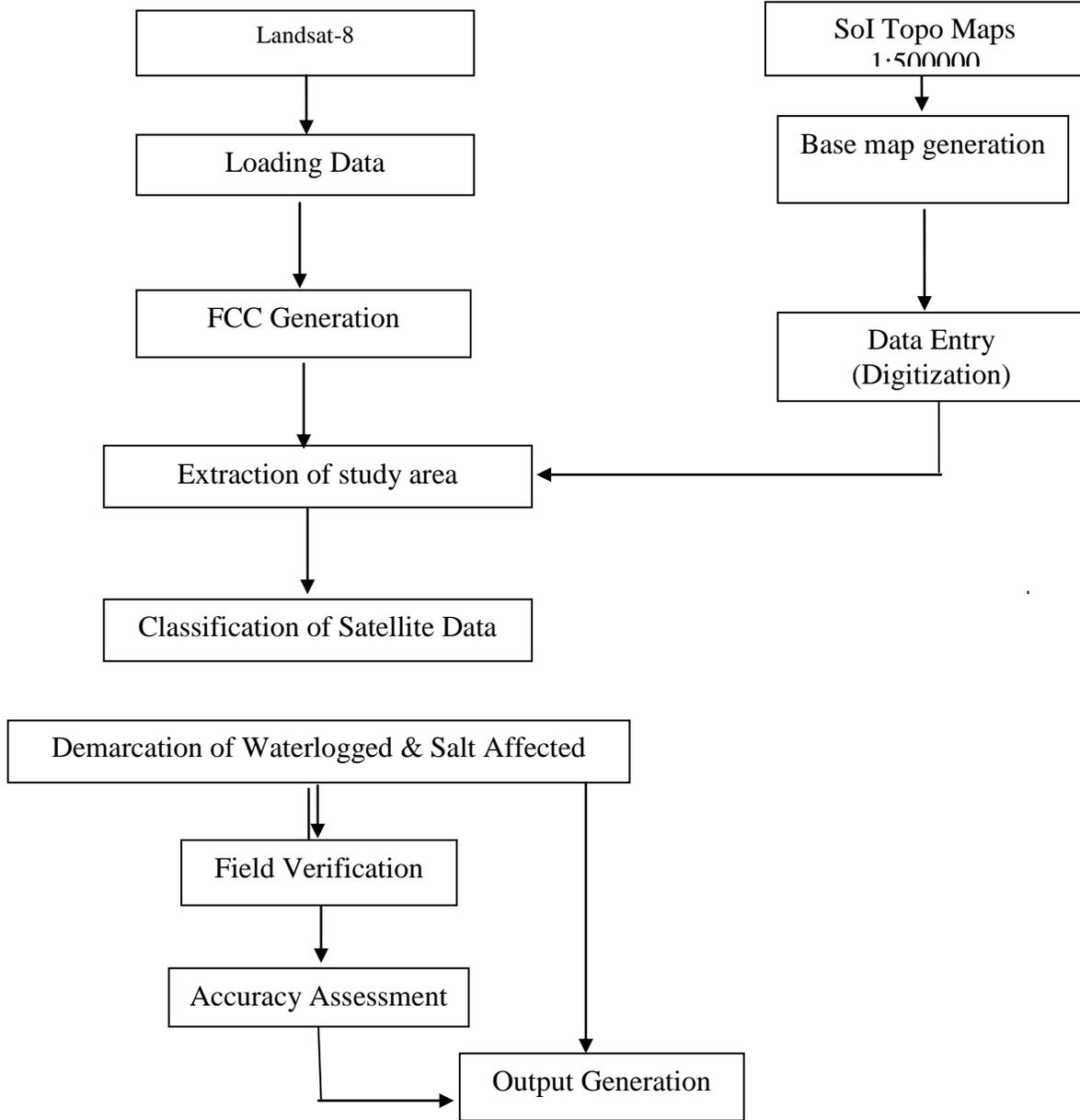


Fig.1 Location map of study area

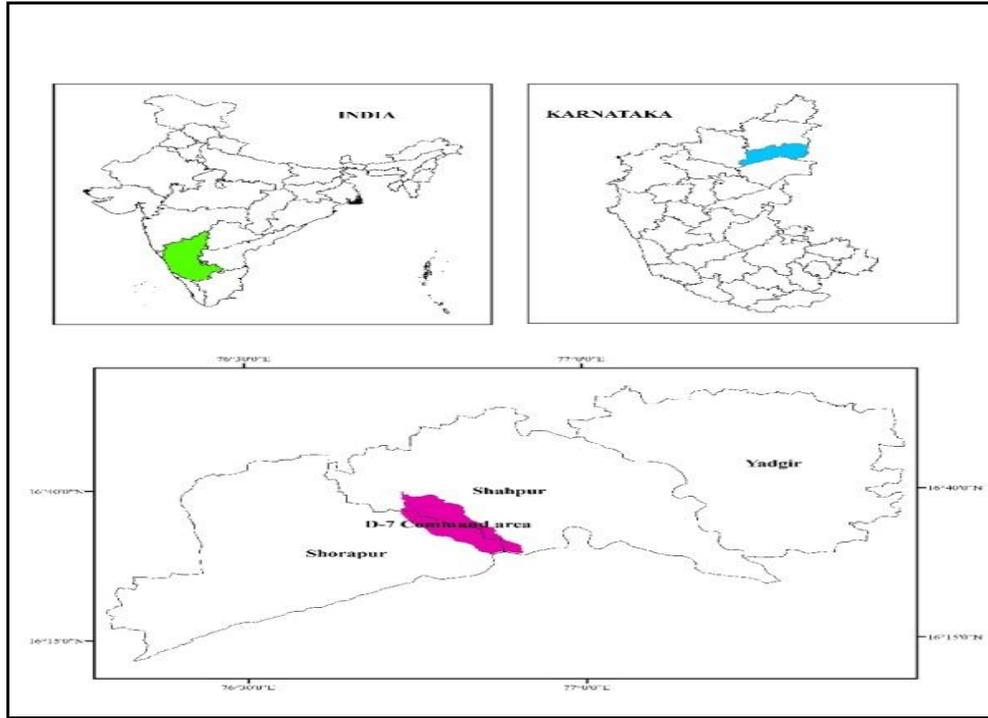
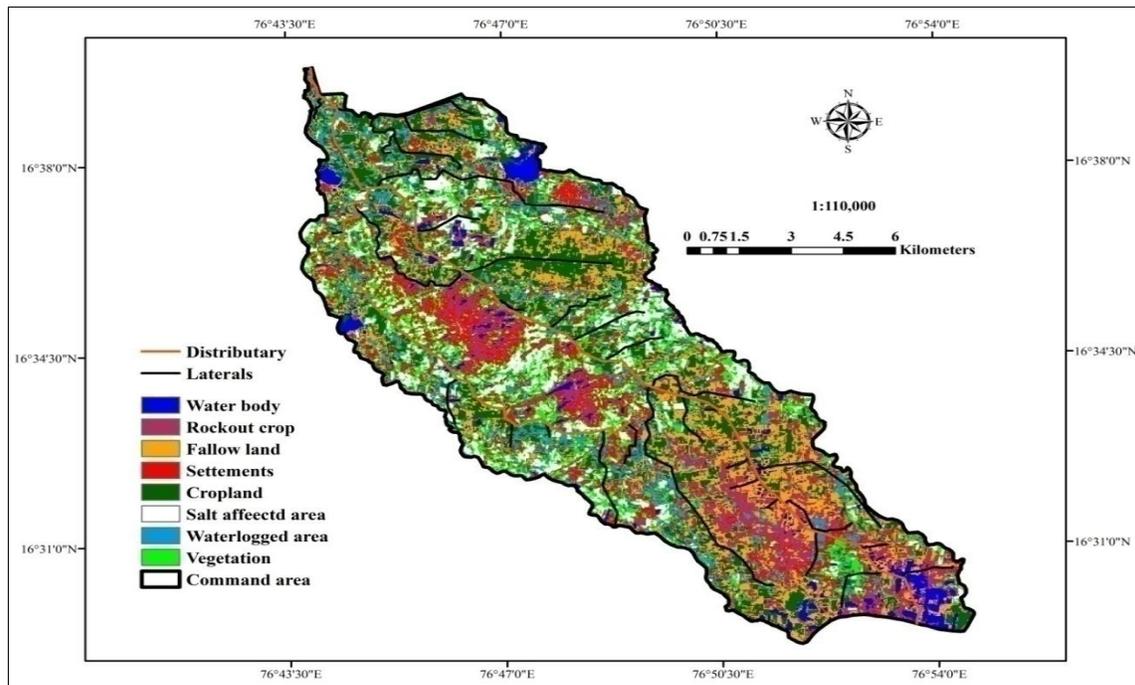


Fig.3 Unsupervised classifications for November-2015 image



Similarly Table 7 indicates accuracy of land use/land cover classification for May 2016. To classify land use/land cover, water body and vegetation has user accuracy 100 per cent. The producer accuracy for water body area was 100 per cent. The least user accuracy was found for salt affected area (71.42 per cent) followed by fallow land (77.77 per cent) and rock out crop (77.77 per cent) and Kappa co-efficient was found 0.81.

Accuracy of land use/land cover classification during May 2017 is presented in Table 8. To classify land use/land cover out of total eight class's, water body, waterlogged area and settlements have user accuracy of 100 per cent. The producer accuracy for water body and waterlogged area was 100 per cent. The least user accuracy of 71.42 per cent was found for vegetation followed by salt affected area 75 per cent and Kappa co-efficient was found 0.84.

In conclusion, unsupervised classification for November 2015 and 2016 cropped area was highest having total area of 2551.77 and 3700.53 ha and minimum area was water body 649.98 and 713.07 ha which accounts for 16.31, 23.65 and 4.15, 4.56 per cent of the cultivable command area respectively during November 2015 and 2016. It was also observed that salt affected area was 1567.44 and 1232.91 ha and waterlogged area was 2033.28 and 1703.07 ha which accounts for 10.02, 7.88 and 12.99, 10.88 per cent of the cultivable command area during November 2015 and 2016 respectively. The comparison of unsupervised classification for May 2016 and 2017 in which it was observed that fallow land area was highest having total area of 2388.87 (May 2016) and 4706.19 ha (May 2017) and minimum area was settlement (1336.86 and vegetation 569.34 ha) which accounts for 15.27, 30.08 and 8.54, 3.64 per cent of the cultivable command area respectively during May 2016 and 2017.

Also, salt affected area was 1558.53 and 1595.52 ha and waterlogged area was 2026.89 and 1149.66 ha which accounts for 9.96, 10.20 and 12.95, 7.35 per cent of the cultivable command area respectively during May 2016 and 2017.

Supervised classification for Nov-2015 and Nov-2016 indicated that cropped area was highest having total area of 6460.62 and 8261.55 ha which accounts 41.28 and 52.79 per cent and minimum area was water body 146.79 and 131.49 ha which accounts for 0.93 and 0.84 per cent of the cultivable command area respectively. During the same period salt affected area was 1797.03 and 1762.11ha which accounts 11.48 and 11.26 per cent and waterlogged area was found 1441.98 and 1489.77 ha which accounts for 9.21 and 9.52 per cent of the cultivable command area respectively. Similarly during May 2016 and 2017 it was observed that fallow land area was highest having total area of 4982.67 and 5815.26 ha which accounts 31.84 and 37.16 per cent and minimum area was waterlogged 52.87 and 53.46 ha which accounts for 0.33 and 0.34 per cent of the cultivable command area respectively. Also the salt affected area was 3486.96 and 3602.86 ha and waterlogged area was found 52.87 and 53.46 ha which accounts for 22.28 and 23.02 per cent of the cultivable command area during May 2016 and 2017 respectively.

References

- Alrababah, M. K. and Alhamad, M. N., 2006, Land use/cover classification of arid and semi-arid Mediterranean landscapes using Landsat ETM. *Int. J. of Remote Sensing.*, 27 (13), 2703-2718.
- Bahadur, K.C.K., 2009, Improving Landsat and IRS image classification: Evaluation of unsupervised and supervised classification through band

- ratios and DEM in a mountainous landscape in Nepal. *Remote Sensing.*, 1 (4), 1257-1272.
- Borghuis, A. M., Chang, K. and Lee, H.Y., 2007, Comparison of automated and manual mapping of typhoon-triggered landslides from SPOT-5 imagery. *Int. J. of Remote Sensing.*, 28 (8), 1843-1856.
- Giles, M. F., 2008, Harshness in image classification accuracy assessment. *Int. J. Remote Sensing.*, 29(11): 3137-3158.
- Megersa, O. D., 2012, Analysing decadal land use/land cover dynamics of the lake Basaka catchment (Main Ethiopian Rift) using LANDSAT imagery and GIS. *Lakes and Reservoirs Rese. Manage.*, 17: 11-24.
- Mukherjee, S., and Mukherjee, P., 2009, Assessment and comparison of classification techniques for forest inventory estimation: A case study using IRS-ID imagery. *Int. J. of Geoinformatics.*, 5 (2), 63-73.
- Thomson, A. G., Fuller, R. M. and Eastwood, J. A., 1998, Supervised versus unsupervised methods for classification of coasts and river corridors from airborne remote sensing. *Int. J. of Remote Sensing.*, 19 (17), 3423-3431.
- Trisurat, Y., Eiumnoh, A., Murai, S., Hussain, M.Z. and Shrestha, R.P., 2000. Improvement of tropical vegetation mapping using a remote sensing technique: a case of Khao Yai National Park, Thailand. *Int. J. of Remote Sensing.*, 21 (10), 2031-2042.

How to cite this article:

Ayyanna, B.S. Polisgowdar, M.S. Ayyanagowdar, Anilkumar. T. Dandekar, G.S. Yadahalli and Bellakki, M.A. 2018. Accuracy Assessment of Supervised and Unsupervised Classification using Landsat-8 Imagery of D-7 Shahapur Branch Canal of UKP Command Area Karnataka. *Int.J.Curr.Microbiol.App.Sci.* 7(07): 205-216. doi: <https://doi.org/10.20546/ijemas.2018.707.025>