LAND USE LAND COVER (LULC) CLASSIFICATION USING FUSION OF MULTI-SENSOR OPTICAL and MICROWAVE DATA FROM SENTINEL MISSION

Vivek Kumar Dr. A.P.J.Abdul Kalam Technical University, Lucknow ,India (vgupta491@gmail.com) , **Ashutosh Bhardwaj** Indian Institute of Remote Sensing, Dehradun, India (ashutosh@iirs.gov.in), **A.L. Haldar** Remote Sensing Application Centre U.P., India (amritlalhaldar@gmail.com)

1 KEY WORDS

Multi-sensor fusion, LULC, Sentinel Mission, SAR, Ehlers Fusion

2 ABSTRACT

Land use land cover mapping from satellite imagery is of great important since it allows to analyze terrain features and is also useful for monitor temporal changes (change detection) like dynamics of water resource, forest cover or urban environment economically. Optical satellite sensors usually detect reflection from features of earth in the visible and infrared part of the electromagnetic spectrum. In contrast, Synthetic Aperture Radar (SAR) has ability to penetrate cloud and also independence of requirement of daylight. These advantages make SAR remote sensing technique different and attractive data source for land-use land-cover mapping. In this study the objective is the mapping of Land use land cover using both Synthetic Aperture Radar (SAR) & Optical remotely sensed datasets through fusion techniques.Sentinel-1A and sentinel-2A mission datasets are used here for parts of Uttar Pradesh .The fusion of data has been done with help of HPF (High pass Filter) resolution merge and Ehlers fusion methods. In addition both the sentinel imageries from the sentinel mission has proved to be the best contemporary possible open source images which suit the study. The fusion of SAR data (sentinel 1A) and multispectral image (sentinel 2A) has improved the separability of classes and hence the accuracy is improved from 76.17% in sentinel 2A image to82.42% and 82.81% in the HPF and Ehlers fused images respectively. Thus, the study depicted that the resultant fused image from multisensory data has better quality for classification.

3 INTRODUCTION

The available open source Optical (multispectral) image is generally used for land cover mapping having different spatial and spectral resolution. Optical satellite sensors usually detected reflection from earth objects in the visible and infrared part of the electromagnetic spectrum .It shows information dependency on reflected light which is the limitation of optical satellites. The optical sensor are used the visible portion of electromagnetic spectrum which cannot penetrate clouds, water vapour and other atmospheric hindrances during inclement weather conditions due to its short wavelength (Aggarwal,2004) thus there must be the lost of information in satellite image. In contrast, Synthetic Aperture Radar (SAR) has ability to penetrate cloud and also independence of requirement of daylight (Mccandless, 1978). These advantages make SAR remote sensing technique different and most attractive data source for land-use/land-cover mapping, especially when optical sensor is not capable to give information due to some its limitations. SAR sensor use the backscattered energy of the different targets which depend on the properties the surface, such as slope, roughness, humidity, textural in homogeneities and dielectric constant. These dependencies allow SAR imagery to be used to separate among various objects on the earth's surface, such as urban area, vegetation and topography (Sciences,2015). Along with these advantages, SAR imagery also has some limitations like it is only appeared in gray scale color thus it is more difficult for visualization and interpretation.

The pre-processing of SAR image is required some necessary tedious work before using to extract proper information.

There are some advantages and limitations in both techniques ,In optical imagery spectral response of some objects like dry river and settlement in the visible and infrared wavelength parts are similar but in SAR data, the responses is very sensitive to varied terrain, roughness and structure, moisture content, for example, settlement showed very high backscattering coefficient value but low value represented by dry river, however, to separate between agriculture and forest is a difficult mission in SAR but can do easily in multispectral images (Nguyen,2016).   In present study the objective is the mapping of land cover using both Synthetic Aperture Radar (SAR) & Optical remotely sensed datasets through fusion techniques. The land covers supervised   classification has been done using SAR and Optical datasets. Both the data (Sentinel-1A and Sentinel-2A) used in this research are open source. This is the attempt to explore the potential and generate a good quality classified Land use land cover map.The study area has been classified into six LULC classes named agricultural land ,River (water body) ,Dry river bed (sand),Built-up, fallow land and forest areas .

4 OBJECTIVE

The objective of this study is to perform a Land use land cover classification using fusion of multi-sensor optical and microwave data from Sentinel mission for the area of Kanpur, Unnao and Lucknow region of Uttar Pradesh, India.

5 STUDY AREA AND DATA SETS

For the present study, an area in between Kanpur and Unnao is taken for study, which lies between 26°12’ and 26°49’ North latitude and 80°12’ and 80°59’ East longitudes with an area of 5382 km² approximately in middle of Uttar Pradesh, India. This area covers District Kanpur, Unnao and some portion of Lucknow.

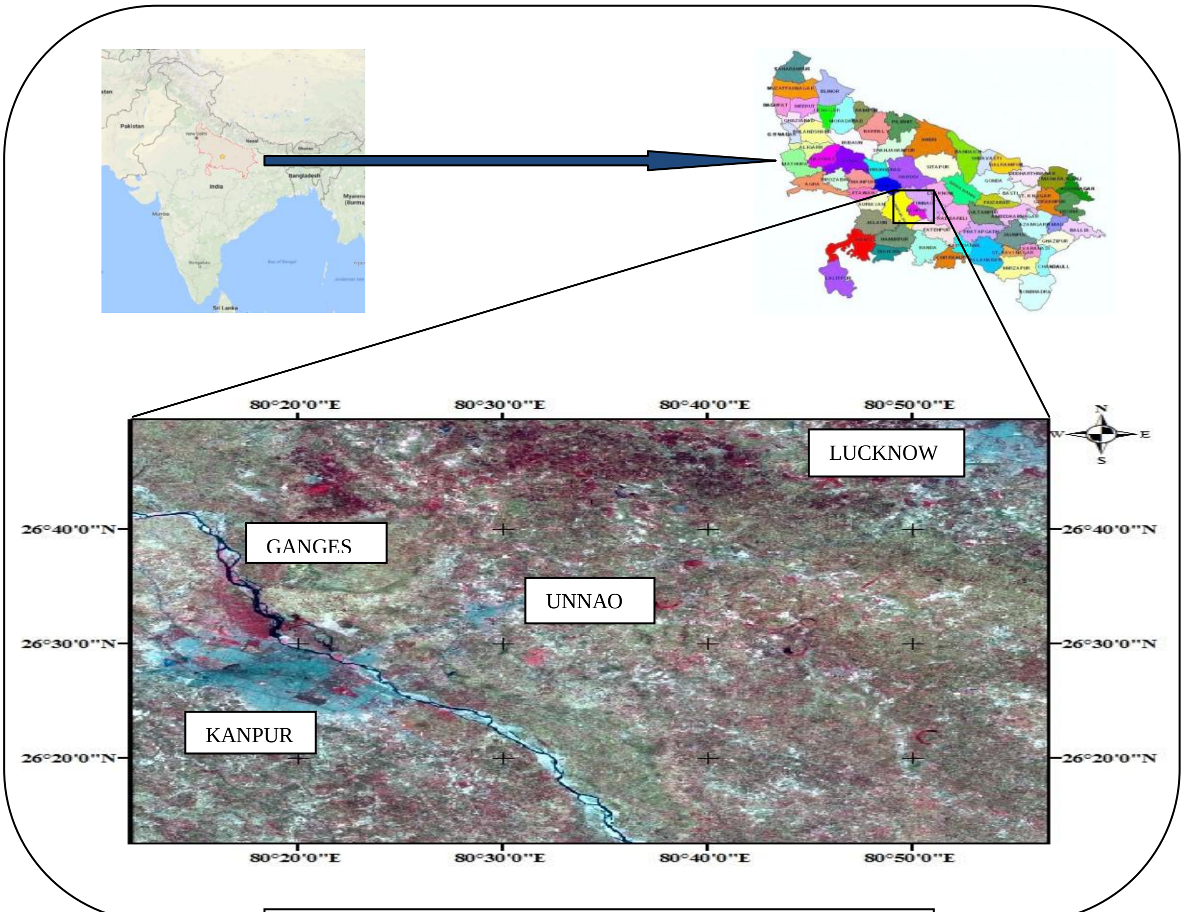

Figure 1: location of study area in image (Sentinel 2A)

It is the central Ganga alluvial plain region having majorly alluvial and sandy soil. The average altitude of Kanpur is 126 m above mean Sea level.

5.1 Sentinel 1A

Sentinel 1A is the space borne synthetic aperture radar (SAR) satellite. The prime objectives of sentinel -1A mission was to monitor Land and Ocean and provide C-Band SAR data continuously after the retirement of ERS-2(European Remote Sensing) satellite. It carries a C-SAR sensor with frequency 5.405 GHz, which offers medium and high resolution imagery in all weather conditions. It is capable of obtaining night imagery and detecting tiny movement on the ground.

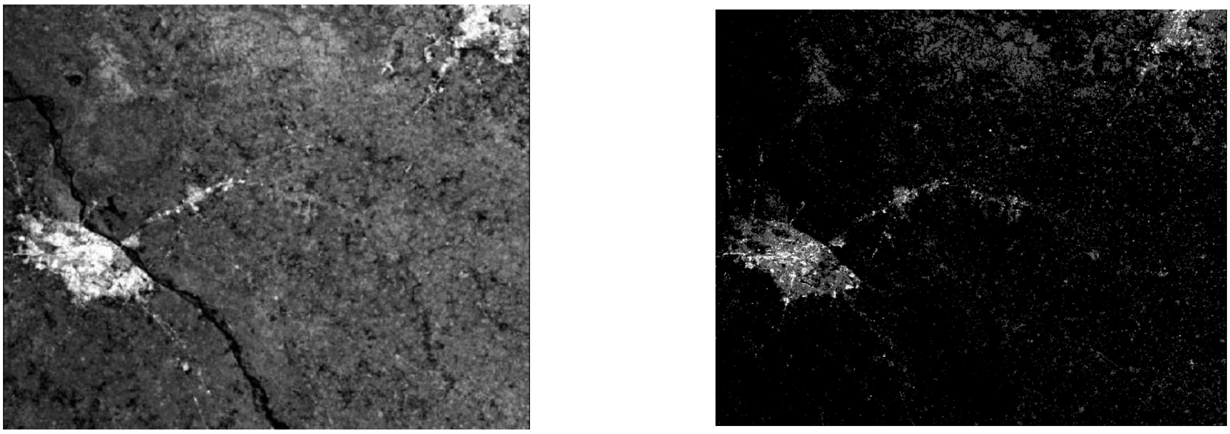

Figure 2: Sentinel-1A Images of the Study Area (VH band-left; VV band-right)

Table 1: Description of Sentinel-1A imagery

Product	**Data Detail**
Instrument name	Synthetic Aperture Radar (C-band)
Instrument mode	IW
Pass Direction	DECENDING
Date	04 march.2016
Time	06:00:03.579006
Orbit Number	10212
Polarisation	VV; VH
Resolution	10m
Product level	L1
Product type	GRD

5.2 SENTINEL 2A

The Sentinel-2 mission is a land monitoring satellite which provides a high resolution optical imagery and provides continuity for the current SPOT and Landsat missions. It contains two satellites in its constellation. It provides a global coverage of the Earth's land surface at every 10 days with one satellite and 5 days with 2 satellites, making the data of great use in recent studies. The satellites are equipped with the state-of-the-art MSI (Multispectral Imager) instrument that offers high-resolution optical imagery (10m, 20m and 60m). The details of Sentinel-2A image used in this study are given in table 2

Table 2: Description of Sentinel 2A imagery

Product	**Data Detail**
Instrument name	Multi-Spectral Instrument (MSI)
Pass Direction	DESCENDING
Orbit number	19
Processing level	Level 1C
Date	04 April 2017
Time	05:06:51.026Z
Spatial Resolution (m)	10;20;60
Cloud cover %	0%

6 METHODOLOGY

The Prime focus of this study is to explore the potential and generate a good quality classified Land cover map using SAR and Optical datasets. Both the data (Sentinel-1A and Sentinel-2A) used in this study are open source datasets and available at European Space Agency (ESA) sentinel DATA HUB. The optical data is available in 12 bands for this study purpose only three bands are used. These three band combination is the standard FCC (false color composite) are B8: NIR (842nm), B4: RED (665nm), B3: GREEN (560nm).The multispectral image is available in Universal Transverse Mercator (UTM) projection system and the SAR image is has default Geographic projection system thus for fusion process optical image transformed into geographic projection system and then the registration of image has been done with the help of geometric correction tool in raster toolbox in ERDAS. Data fusion combines data from many sources or time periods in an attempt to improve the information, leading to a better result in proceeding actions, such as a classification (Hall, 2004).Data fusion can be performed on three different processing levels, defined by the stage at which the fusion is done. These stages are the pixel level, the feature level and the decision level (Pohl, 1998). There are many fusion techniques are available but for land cover mapping only those techniques are required which give better result .On the basis of visual interpretation and edge detection of the fused image Ehlers fusion and High Pass Filter (HPF) fusion images has been considered for classification.

The supervised classification of both fused images and Sentinel 2A (standard FCC) image has been done for further comparison. Supervised classification is the process of giving samples of known identity (training area) to the classify pixel with unknown identity. Pixels covered within this area are known as training samples. These training samples guide to the classification algorithm to assigning specific spectral values to given informational class. The process has been performed in ERDAS IMAGINE 2014.

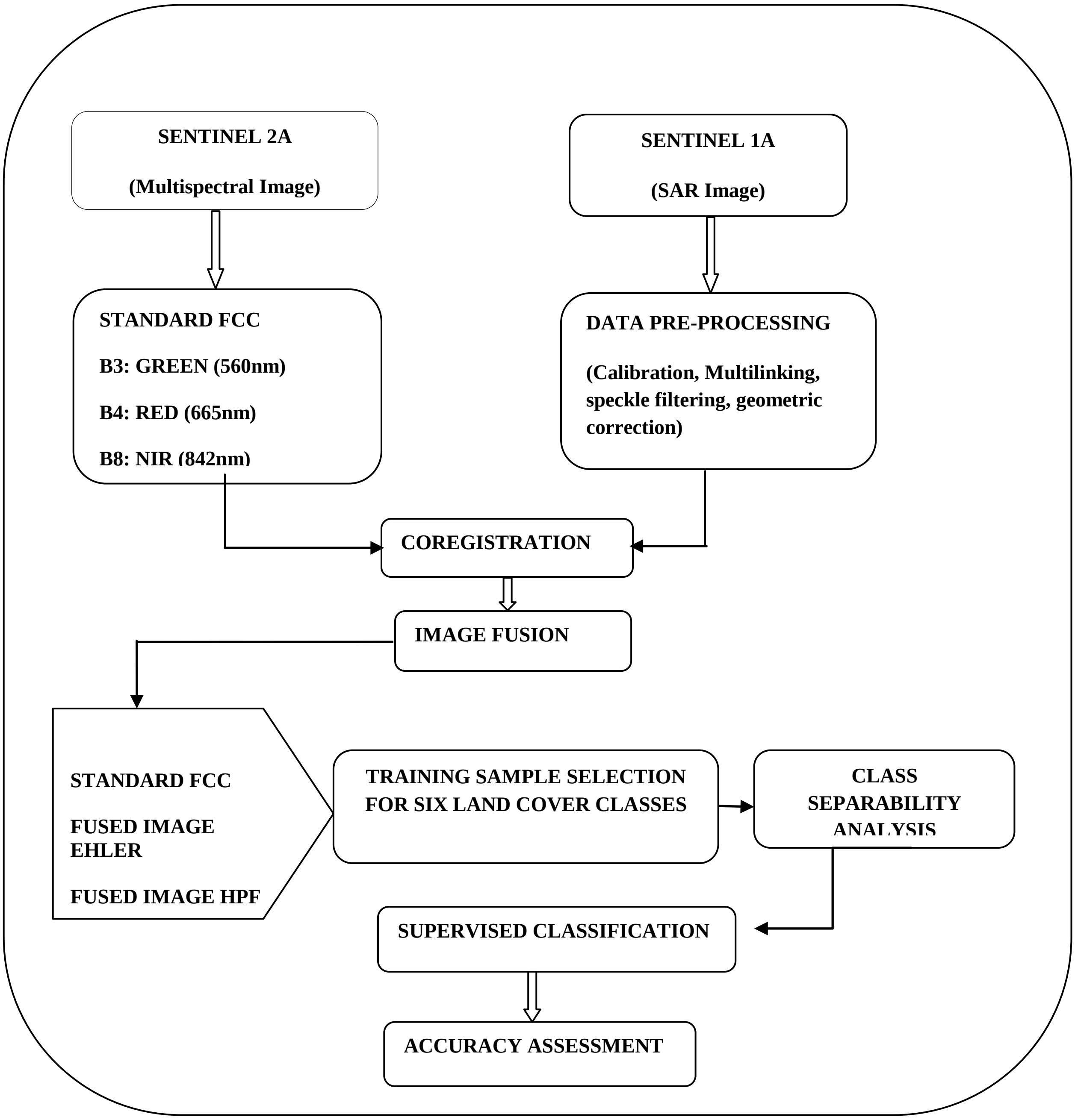

Figure3: Flowchart for LU/LC Classification

6.1 Pre-processing of Data

To process and extract information from SAR data, pre-processing is necessary that removes radiometric and geometric distortions which is unique to SAR data. In pre-processing the radiometric calibration, multi-looking, speckle filtering and geometric correction (geocoding) performed in sequence. This process has been done in The Sentinel Application Platform (SNAP) toolbox which is also available at ESA official web portal. After SAR data pre-processing, subset of study area clipped and export in Geotiff format. The basic steps of supervised classification have been followed here.

6.2 Training Data

Training cells are the area of known identity delineated on the given satellite image. It is usually done by providing the corners of rectangular polygonal area. The specific training Ares are unique for each land cover classes which is to identify a set of pixels that represents the spectral variation within each class. The accuracy of classification is depending on the quality of training areas. The study area has been divided into six classes namely built-up, agriculture, forest, and dry sand, fallow land and water body. The training sample was taken from both fused image and from standard (8-4-3) sentinel 2A image for determining the classes. After taking the training samples class signature separability analysis performed. It shows the separation between two classes in provided training samples. There are four methods in ERDAS for computing signature separability between training samples .These methods are Euclidean distance, Divergence, Transformed Divergence and Jefferies- Matusita Distance method. The transformed divergence method is used in this study because it is the most efficient method among the four techniques (Duadze, 2004).

Table 3: Separability analysis for various land cover classes in different images

CLASSES	Sentinel-2A (NIR-R-G)	HPF FUSED	EHLER FUSED
Forest - Fallow	1992	1997	2000
Agriculture-Fallow	1251	1970	1975
Built up - Dry River sand	669	1990	2000
Built up - Fallow	1437	1705	1964
Fallow- Dry river sand	1413	1930	1855
Forest - Agriculture	1960	1971	1959

The scale of divergence value has range between 0 to 2000 (Jenson, 1996).As a general rule ,if separation value is greater than 1900 ,then the classes can be separable .when class separation value is between 1700 and 1900 ,the separation is fairly good, below 1700 ,the separation is poor. From the above table it can be seen easily that the class separability is enhanced in fused images. as in multispectral image the built-up class and fallow land class is poorly separable (1437)and is in EHLER fused image it is separable(1964) and fairly separable in HPF fused image(1705) .The analysis indicates the spectral quality improvement in fused image with reference to initial multispectral image.

6.3 Selection of appropriate classification algorithm

There are various supervised classification algorithms are available in ERDAS like Maximum likelihood, Mahalanobis distance, Minimum distance, Spectral angle mapper ,Spectral correlation mapper. The choice of a particular classifier or decision rule depends on the nature of the input data and desired output. The maximum likelihood algorithm is used in present study. This algorithm assigns each pixel having pattern measurements or features X to the particular classes whose units are most probable or likely to have given rise to feature vector X. It assumes that the training data statistics for each class in each band are normally distributed (Lillesand, 1993) Maximum likelihood Classification is a statistical decision criterion to assist in the classification of overlapping signatures; pixels are assigned to the class of highest probability.

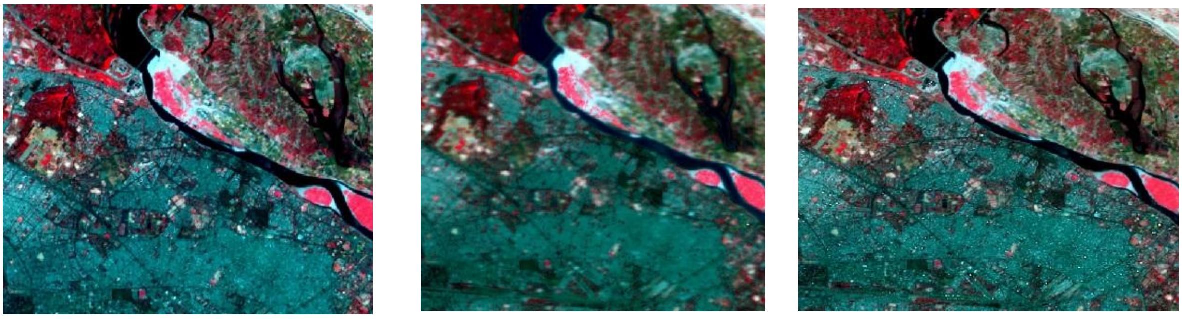

Figure 4: Subsets of (i) sentinel 2A (ii) Ehler and (iii) HPF fused image

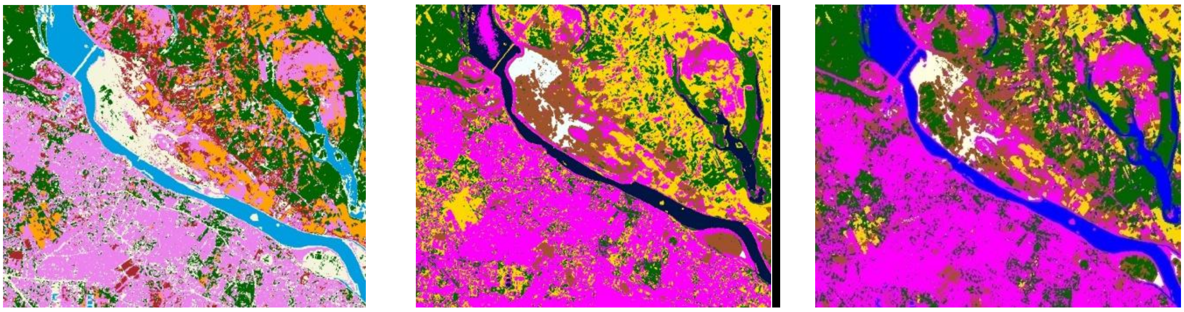

Figure 5: classified images of i) Sentinel 2A ii) Ehler and iii) HPF fused image

6.4 Accuracy Assessment

There is a requirement of insitu information or previous knowledge of ground truth for assessment of classified image accuracy .This information can be collected through Google earth which can be compared with the derived classified map. One of the most common ways to express the classification accuracy is the preparation of classification error matrix also known as contingency matrix or confusion matrix. For this assessment the 256 points are taken by the software that cover full image .user have to check one by one all selection and put give the correct value of the class on the basis of ground truth . For assigning these random points value we can take the help of Google earth. The ERDAS calculate the accuracy according to these assigning values .

6.4.1 Classification error matrix

Error matrix comparison is done on the basis of class by class division. The relationship between known referenced data (ground truth) and comparison is on a category by category basis. The relationship between known referenced data (ground truth) and the corresponding result of classification. Such matrix is square matrices. From this error matrix the study of the various classification errors of exclusion and inclusion is become easier. The training data pixels that are classified into their proper land cover categories are located in the major diagonal axis of this matrix (from upper left to lower right).all non-diagonal elements of the matrix represents the error of omission (exclusion) or commission (inclusion) (M. Sandberg, 2016). The accuracy of a particular category can be calculated by dividing the number of correctly classify pixels to the total number of pixel in corresponding rows or column according to find user accuracy or producer accuracy respectively.

Table 4: Confusion matrix of sentinel-2A classified Image

Classified data	Built-up	Water body	Forest	Dry sand	Fallow land	Agriculture	Total row
Built-up	55234	0	98	57408	1842	445	115027
Water body	0	7187	0	7187	0	0	14374
Forest	923	3	33548	1723	1209	1187	38593
Dry sand	5914	165	46	21421	1349	130	29025
Fallow land	6089	2	1046	8999	46313	5411	67860
agriculture	520	0	911	1216	7539	47969	58155
Total column	68680	7357	35649	97954	58252	55142	323034

Table 5: values of accuracy of different classes

Classes	Producers accuracy	User accuracy	Kappa coefficient
Built-up	41.67%	76.92%	0.7454
Water body	16.67%	100%	1.00
Forest	76.32%	70.73%	0.6563
Dry sand	21.86%	73.80%	0.00
Fallow land	71.67%	66.15%	0.5579
Agriculture	8 7.90	81.95%	0.6500

The Overall classification accuracy is 76.17% and Overall Kappa Statistics is 0.6390.As the producers accuracy is varied from 16.67% of water body to 87.90% of agricultural land and user accuracy is varied from 66.15 % to 100 % .but the overall accuracy is good. The error matrix is depends upon training data. If the results are good it means the training sample is spectrally separable and good classification is done in training sample areas. The same classification analysis is performed over HPF fused image and Ehler based fused image .the error matrix and accuracy assessment report is given in the following tables.

Table 6: confusion matrix of HPF fused image classification

Classified data	Built-up	Fallow land	Water body	Forest	Dry sand	Agricultur e	Total row
Built-up	85235	343	12	281	91	41	86003
Fallow land	4705	22563	0	628	317	508	28721
Water body	106	0	13165	0	0	0	13271
Forest	1482	162	472	38355	0	172	40643
Dry sand	378	163	0	0	9587	2	10130
agriculture	32	138	0	1118	0	14667	15955
Total column	91938	23369	13649	40382	9995	15390	194723

Table 7: Values of accuracy of different classes

Classes	Producers accuracy	User accuracy	Kappa coefficient
Dry sand	66.67%	100%	1
Forest	87.10%	62.79%	0.5766
Built-up	69.23%	81.82%	0.7976
Waterbody	20.00%	100%	1
Agriculture	84.55%	93.94%	1.8937
Fallow land	87.5%	78.65%	0.6895

The Overall Kappa Statistics is 0.7474 and Overall Classification Accuracy is 82.42%.

Table 8: Confusion matrix of Ehlers fused image classification

Classified data	Built-up	Forest	Water body	Dry sand	Fallow land	Agriculture	Total row
Built-up	28974	8	6	5	337	440	29770
Forest	116	22507	12	0	2	903	23540
Water body	0	0	3480	0	0	0	3480
Dry sand	0	0	0	17618	394	0	18012
Fallow land	419	1	0	583	30165	1210	32378
Agriculture	311	597	2	0	73	44988	45971
Total column	29820	23113	3500	18206	30971	47541	153151

Table 9: Values of accuracy of different classes

Classes	Producers accuracy	User accuracy	Kappa coefficient
Built-up	55.17%	80.00%	0.7744
Waterbody	37.50%	100%	1
Forest	90.32%	84.85%	0.8276
Dry sand	-	-	0
Fallow land	77.78%	85.14%	0.7825
Agriculture	97.14%	80.95%	0.6771

The Overall Kappa Statistics is 0.7479 and Overall Classification Accuracy is 82.81%.

7 CONCLUSION

This study investigated the potential of Open source optical satellite data and the suitability of fusion techniques for land use/land cover mapping. Classification was performed using maximum likelihood algorithm. Overall classification accuracies of about 82.81% could be reached in classifications distinguishing six major classes named agricultural land ,River (water body) ,Dry river bed (sand),Built-up, fallow land and forest areas . The fusion of multispectral (Sentinel 2A) and SAR (Sentinel 1A) image has been done with help of HPF (High pass Filter) and Ehlers fusion methods .It reflected the improvement in the separability of classes. The accuracy has been improved from76.17% in sentinel 2A to 82.42% and 82.81% in the HPF and Ehlers fused image respectively. The overall results and outcome of the study strengthen the application of Fusion of SAR and multispectral imagery for LULC classification. One major contribution of this study is the establishment of a method that enables the user to improve the quality of freely available satellite data. The insights and promising results from this study will hopefully contribute to continued research in this direction.

8 ACKNOWLEDMENT

The authors wish to thank Dr.A.Senthil Kumar, Director IIRS Dehradun and Director, Remote sensing Applications Centre, U. P. for providing a conducive environment and facilities for study and research. Authors are also thankful to the open source data and tool providers such as United States Geological Survey (USGS) and European Space Agency (ESA).

9 REFERENCES

[1] B. Sciences, “Soil Moisture Retrieval Using Synthetic Aperture Radar (SAR) data,” no. January, 2015.
[2] ERDAS IMAGINE 8.4 Field Guide: ERDAS Inc.
[3] Hall, D.L. & McMullen, S.A.H. 2004. Mathematical Techniques in Multisensor Data Fusion. 2^{nd}

ed. Norwood, MA: Artech House. ISBN 9781580533355.
[4] Jensen, J.R., 1996. Introductory digital image processing: a remote sensing perspective. Prentice-Hall, Upper Saddle River, NJ, 318 p.
[5] Lillesand, T.M. and Kiefer, R. 1993. Remote Sensing Image Interpretation. John Wiley, New York.
[8] "Manh-Van-Nguyen_Land-Cover-Mapping-Using-Integration-Of-Sar-And-Optical-Remotely-Sensed Datasets."

[9] M. Sandberg, "Land cover mapping with multi-temporal SAR and optical satellite data," no. February, 2016.

[10] Pohl, C. & Van Genderen, J.L. 1998. "Multisensor image fusion in remote sensing: Concepts, methods and applications". International Journal of Remote Sensing, vol. 19, no. 5, pp. 823-854. ISSN 0143-1161.

[11] R. Bamler, "The SRTM mission: A world-wide 30m resolution DEM from SAR interferometry in 11 days," Photogramm. Week, pp. 145–154, 1999.
[12] S. Aggarwal, "Principles of remote sensing," Satell. Remote Sens. GIS Appl. Agric. Meteorol., pp. 23–38 , 2004.
[13] Stephen Edem Korbla *Duadze* ... Bonn, Univ., Diss., *2004* ISBN 3-86537-041-1 1. Referent ... Dr. Paul L.G. Vlek Tag der Promotion: 19.01.*2004*
[14] S. W. W. Mccandless and C. R. Jackson, "Chapter 1 . Principles of Synthetic Aperture Radar," SAR Mar. User's Man., pp. 1–23, 1978.
[15] Z. A. Keninger, "Assessing land cover map accuracy and performance of hydrological models for small stream catchments using GIS," 2012.