COMPARISON OF SUPERVISED CLASSIFICATION TECHNIQUES WITH ALOS PALSAR SENSOR FOR ROORKEE REGION OF UTTARAKHAND, INDIA

Amit Kumar Shakya [1, *], Ayushman Ramola [1], Akhilesh Kandwal [1], Rishi Prakash [1]

[1] Department of Electronics and Communication Engineering, Graphic Era, Dehradun, India - xlamitshakya.gate2014@ieee.org

Commission V, SS: Emerging Trends in Remote Sensing

KEYWORDS: Advanced land observing satellite, PALSAR, Supervised classification, Kappa coefficient

ABSTRACT

The Advanced Land Observing Satellite (ALOS) is developed by the Japanese Aerospace Exploration Agency (JAXA) which was launched in the year 2006 for the Earth observation and exploration purpose. The ALOS was carrying PRISM, AVNIR-2 and PALSAR sensors for this purpose. PALSAR is L-Band synthetic aperture radar (SAR). The PALSAR sensor is designed in a way that it can work in all weather conditions with a resolution of 10 meters. In this research work we have made an investigation on the accuracy obtained from the various supervised classification techniques. We have compared the accuracy obtained by classifying the ALOS PALSAR data of the Roorkee region of Uttarakhand, India. The training ROI'S (Region of Interest) are created manually with the assistance of ArcGIS Earth and for the testing purpose, we have used the Global positioning system (GPS) coordinates of the region. Supervised classification techniques included in this comparison are Parallelepiped classification (PC), Minimum distance classification (MDC), Mahalanobis distance classification (MaDC), Maximum likelihood classification (MLC), Spectral angle mapper (SAM), Spectral information divergence (SID) and Support vector machine (SVM). Later, through the post classification confusion matrix accuracy assessment test is performed and the corresponding value of the kappa coefficient is obtained. In the result, we have concluded MDC as best in term of overall accuracy with 82.3634% and MLC with a kappa value of 0.7591. Finally, a peculiar relationship is developed in between classification accuracy and kappa coefficient.

1. INTRODUCTION

Image classification into several categories or classes followed by an assessment of classification accuracy is complex as well as an interesting area of research in the field of remote sensing (RS) [1]. The classification of the RS data is performed to identify a particular area of interest or class, followed by estimation of the accuracy of the classification [1]. Researchers are continuously working to develop a quick and highly accurate algorithm for image classification. Image classification techniques are basically classified into two categories i.e. supervised and unsupervised classification, in the supervised classification ground truth observations regarding the class of data are known already to the user [2]. In unsupervised algorithms, the user is not aware of the initial conditions of the data, field or the area of observation. The supervised classification techniques are categorized as parallelepiped [3] minimum distance [4], mahalanobis distance [5], , maximum likelihood [6], spectral angle mapper [7], spectral information divergence, [8], binary encoding [9], artificial neural network, [10] and support vector machine [11]. Unsupervised classification techniques include Isodata, [12] and k-means algorithm [13]. H. Zhuang [14] developed an approach based on the combination of SAM and change vector analysis (CVA) for the unsupervised classification of the remote sensing satellite multispectral data. E. Hasan [15] developed a model based on the combination of the SAM and surface structure of the Central Eastern Egyptian desert to map gold mines sites. D.Renza [16] developed a new unsupervised change detection technique based upon the differencing of multi-temporal multi-spectral images. The comparison of the images is obtained by applying SAM between each multi-temporal image and the reference spectrum. E. Zhang [17] improved the classification performance of the hyperspectral images using a sparse representation classifier based on SID. M. Khaleghi [18] used SID, SAM and principal component analysis (PCA) change detection technique to enhance the alteration of the Sarduiyeh area of Kerman province of Iran. M. Janati [19] used the fusion of SID, SAM and Advanced Spaceborne Thermal Emission and Reflection

Radiometer (ASTER) data for the geographical mapping of the basement domain in Arid Regions. H.Jiao [20] used a combination of the SID and SAM supervised classification techniques to develop a novel algorithm for the computer-based spectral encoding. M.Musci [21] used binary encoding (BE) and SAM technique to compute the local binary pattern (LBP), variance, LBP + variance and LBP/variance for the texture classification of the remotely sensed satellite images. G.Sheng [22] used BE and SAM technique for the feature extraction of the coding based high-resolution satellite scene classification. Y.Shao [23] compared SVM and Classification And Regression Trees (CART) algorithm on the MODIS time series dataset the accuracy of the classification obtained in this case is 91% and a kappa value of 0.79. D.T.Bui [24] compared the SVM, neural network, kernel logistic regression and logistic model tree for the spatial model prediction of the landslide areas. M.B.Kia [25] used the neural network model with Geographical Information System (GIS) for the flood detection across Johar river basin, Malaysia. V.F.R.Galiano [26] used the SVM and neural network classification technique to evaluate the effectiveness of the random forest classifier for the land cover classification. U. Maulik [27] developed a technique of land use/land cover remote sensing images classification through SVM. H.Hong et.al [28] investigated 445 landslide regions, they used SVM and random subspace model to create a new model named as 'random subspace-based support vector machines (RSSVM)' to make an investigation on various landslides activities. X.Wang [29] used the SVM technique to develop a novel WSL model adding human gaze annotation for the classification of the satellite images. In this research, we have investigated the relation in between the kappa coefficient and the classification accuracy when the classification is made through various supervised classification techniques. The comparison is made in between PC, MDC, MaDC, MLC, SAM, SVM and SID techniques. Finally for the estimation of the accuracy post classification confusion matrix and a kappa coefficient of the classified data is created. The confusion matrix consists of user accuracy [30], producer accuracy [30], omission error [30] and commission error [30] which are computed for the individual class of the classified image, finally, the overall accuracy

---

* Corresponding author

of the classified data is computed. Kappa coefficient [*31*] is calculated for the various classified images which is the degree of agreement among various raters. Now from the image matrix shown in fig. 1 procedure to calculate user accuracy, producer accuracy, omission error, commission error and overall accuracy is shown.

CLASSES	REFERENCE DATA				
	Soil	Urban	Vegetation	Water	Row Total
Soil	28	2	5	9	44
Urban	4	42	4	2	52
Vegetation	7	8	64	9	88
Water	1	2	8	32	43
Column Total	40	54	81	52	227

(The leftmost column "CLASSIFIED DATA" spans all data rows)

Fig. 1 Example for image classification

Overall accuracy is defined as the ratio of a total number of correctly classified pixels to the total number of pixels.

$$Overall\ Accuracy = \frac{28+42+64+32}{227} = 73.12\% \quad (1)$$

Omission error refers to the pixels that are unintentionally omitted from the correct class.

$$Ommission\ error\ (soil) = \frac{4+7+1}{40} = 30\% \quad (2)$$

$$Ommission\ error\ (urban) = \frac{2+8+2}{54} = 22\% \quad (3)$$

$$Ommission\ error\ (vegetation) = \frac{5+4+8}{81} = 20.98\% \quad (4)$$

$$Ommission\ error\ (water) = \frac{9+2+9}{52} = 38.46\% \quad (5)$$

Commission error refers to the pixels of another class that are wrongly included in class under consideration.

$$Commission\ error\ (soil) = \frac{2+5+9}{44} = 36.36\% \quad (6)$$

$$Commission\ error\ (urban) = \frac{4+4+2}{52} = 19.23\% \quad (7)$$

$$Commission\ error\ (vegetation) = \frac{7+8+9}{88} = 27.27\% \quad (8)$$

$$Commission\ error\ (water) = \frac{1+2+8}{43} = 25.58\% \quad (9)$$

User accuracy is the accuracy from the point of view of the 'map user' not from the 'map maker'. User accuracy is defined as follows

$$User\ Accuracy = 100\% - commission\ error \quad (10)$$

$$User\ Accuracy\ (soil) = 100\% - commission\ error\ (soil) = 100 - 36.36 = 63.64\% \quad (11)$$

$$User\ Accuracy\ (urban) = 100\% - commission\ error\ (urban) = 100 - 19.23 = 80.77\% \quad (12)$$

$$User\ Accuracy\ (vegetation) = 100 - commission\ error\ (vegetation) = 100 - 27.2 = 72.73\% \quad (13)$$

$$User\ Accuracy\ (water) = 100\% - commission\ error\ (water) = 100 - 25.58 = 74.42\% \quad (14)$$

Producer accuracy is the accuracy from the point of view of a map maker, this provides us information about how much accurately the class will be actually present on the ground.

$$Producer\ Accuracy = 100\% - omission\ error \quad (15)$$

$$Producer\ Accuracy\ (soil) = 100\% - omission\ error\ (soil) = 100 - 30 = 70\% \quad (16)$$

$$Producer\ Accuracy\ (urban) = 100\% - omission\ error\ (urban) = 100 - 22 = 78\% \quad (17)$$

$$Producer\ Accuracy\ (vegetation) = 100 - omission\ error\ (vegetation) = 100 - 20.98 = 79.02\% \quad (18)$$

$$Producer\ Accuracy\ (water) = 100\% - omission\ error\ (water) = 100 - 38.46 = 61.54\% \quad (19)$$

Finally, the kappa value is generated to rate the accuracy of the classification. Kappa coefficient provides information regarding how well the classification of any dataset is performed [*32*]. The statistical range of the kappa coefficient is [-1, 1]. The mathematical expression for the kappa coefficient is defined as under

$$k = \frac{prob(O) - prob(e)}{1 - prob(e)} = 1 - \frac{1 - prob(o)}{1 - prob(e)} \quad (20)$$

Here $prob(o)$= relative observed agreement among various raters (this is identical to the accuracy).

$prob(e)$= is the hypothetical probability of the chance agreement.

Now for the categories 'k', number of the elements 'N', the number of the time rater 'n_{k_i}', and for the predicted category k then the hypothetical probability of the chance agreement is defined as

$$prob(e) = \frac{1}{N^2} \sum_k n_{k_1} n_{k_2} \quad (21)$$

Following a specific conclusion can be derived from the kappa coefficient.

- If the value of the kappa coefficient lies close to -1 then classification is a worse classification.
- If the value of the kappa coefficient lies closer to zero than it is nothing but a random classification.
- Finally, if the kappa coefficient lies close to 1 then it is significantly better and close to accurate classification.

The value of the obtained kappa coefficient is related with the accuracy of various supervised technique to establish a relationship pattern in between classification accuracy and the kappa coefficients.

2. BACKGROUND OF THE ALOS PALSAR PROGRAM

The Advanced Land Observing Satellite (ALOS) was launched on 24th January 2006. The main objective of this satellite was to monitor changes in the land cover & environment using high- resolution satellite images. It was carrying three strong sensors for human visual interpretation, a high-resolution optical sensor, and an active microwave (L band synthetic aperture radar) sensor. ALOS was working on four mission objectives i.e. disasters observations, resource exploration, regional observations and cartography, but the main objective of ALOS was monitoring the content of water, carbon and changing global climate [*33*]. Three different sensors are named as Advanced Visible and Near Infrared Radiometer 2 (AVNIR-2)

[34], Panchromatic Remote Sensing Instrument for Stereo Mapping (PRISM) [35] and the Phased Array L-band Synthetic Aperture Radar (PALSAR) [36], [37]. PALSAR is L band synthetic aperture radar designed to work under all sought of weather conditions. It is designed to provide data in four polarization combinations (HH, VV, VH and HV), where first H or V represents transmit polarization and second H or V represents receive polarization. Application of the PALSAR data includes sea surface ice monitoring, land use/land cover classification, trees height estimation and interferometry. PALSAR consists of some unique features it is designed to operate in three different modes i.e. fine, scan-SAR and polarimetric its center frequency of operation is 1270 MHz, chirp bandwidth is 14 and 28 MHz, range resolution is 7 to 100 m, incidence angle can vary from 8 to 60 degree, observational swath vary from 40 to 350 km, bit length vary from 3 to 5 bits, data rate vary from 120 Mbps to 240 Mbps. The radiometric accuracy for the images obtained from PALSAR is 1dB/orbit i.e. 1.5 dB.

3. SUPERVISED CLASSIFICATION TECHNIQUES

Supervised classification is the technique mostly used for the classification of the remote sensing image data, it is dependent on the suitable algorithm and procedures to classify or label the pixels into a particular class of interest. In these technique prototypes pixels of each desired class are assigned the name '*training data*,' and on the basis of the training data image is classified into several classes, which is again followed by matching of the classified image through '*testing data*' [38]. Testing data contains the information based upon the ground reality, GPS coordinates, topographic maps, etc. Finally, the accuracy test of the classification is performed to obtain the percentage of correct classification.
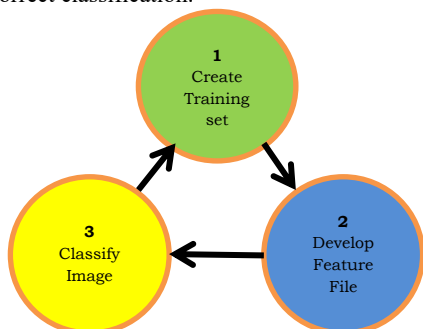

Fig.2 Supervised classification diagram

Now we are discussing about various supervised classification techniques.

Parallelepiped classification: In this classification technique a decision-based rule is used to classify the data. An *n*-dimensional decision boundary is created in the image data space. Here mean of each class is selected which is based on the standard deviation threshold. Pixels values lying below and above the threshold for all *n*-bands are classified. If a pixel value falls under several classes, classifier assigns that pixel to the first matched class, whereas remaining falls under the unclassified category.

Minimum distance classification: In this classification scheme classifier computes Euclidean Distance (ED) between the mean value of the class and the pixel under consideration, and then the pixel is allocated to the class which is at the minimum ED.

Mahalanobis distance classification: This is a direction sensitive based distance classifier, in which statistics derived from each class are used. It assumes that all classes have equal covariance. Here

pixels are classified to the closed ROI class unless a specific threshold is decided else the pixel will be considered as unclassified.

Maximum likelihood classification: In this classification pixel with maximum likelihood is classified to a particular class. It is expressed by the posterior probability of a pixel belonging to the class 'c'.

$$L_c = p\left(\frac{c}{x}\right) = (p(c) * p(\tfrac{x}{c})) / \sum (p(i) * p\left(\frac{x}{i}\right)) \quad (22)$$

Where $p(\frac{x}{c})$ is the conditional probability to analyze x from class c.
x is the image data from the n band.
$p(c)$ is assumed to be equal to $\sum (p(i) * p\left(\frac{x}{i}\right))$.

Spectral angle mapper: This classification technique is based on the assessment of the spectral similarity in the (*n*-dimensions) features.
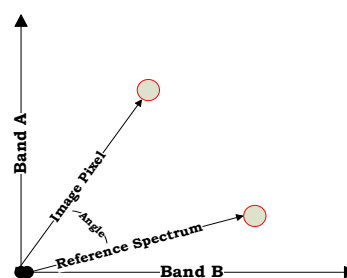

Fig. 3 Supervised spectral angle mapper

This similarity is evaluated by the vector starting from the origin. The length of the vectors is the representation of the reflection intensity. The difference between the two spectra is described by the spectrum angle. Finally, by evaluating the difference between the pixel and the reference spectrum an image can be classified in any number of classes. The angle says α (alpha) formed in between the two vectors is the arc cosine it is calculated on the basis of the the product of all bands $'n \times b'$, total pixels $'t'$, and all the reference pixels denoted by $'rp'$. Then the angle α is represented as follows

$$\alpha = cos^{-1}\left[\frac{\sum_{i=1}^{n \times b} t_i rp_i}{\left(\sum_{i=1}^{n \times b} t_i^2\right)^{1/2}\left(\sum_{i=1}^{n \times b} rp_i^2\right)^{1/2}}\right] \quad (23)$$

Here α = the spectral angle formed between the two vectors
$n \times b$ = total numbers of the spectral bands
t= target pixels
rp =reference pixels

Support Vector Machine: In this classification scheme any dataset is classified in the best hyperplane, which separates data points of one class from another class. Best hyperplane means the one with the largest margin in between classes. The support vector machine represents the data points that are close to the hyperplane. These points lie in the vicinity of the vector slab.
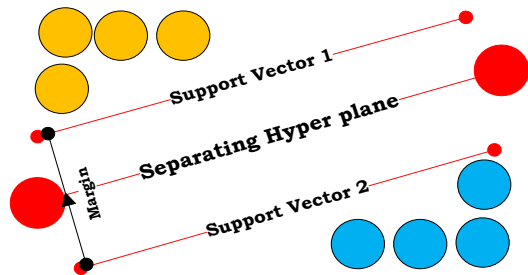
Fig. 4 Supervised support vector machine

Spectral information divergence: In this method divergence measures are used to match pixel with the reference spectrum. Smaller the divergence chances of similarity in the pixels get increased. Pixels which do not meet specified maximum divergence threshold falls under the unclassified category.

4. METHODOLOGY

In this experiment we are performing a comparative accuracy assessment test by comparing the classification accuracy of the PC, MDC, MaDC, MLC, SAM, SVM and SID supervised classification techniques. We have obtained the accuracy of classification and estimated the kappa coefficient value followed by the development of a relationship between overall accuracy and kappa coefficient.

4.1 Geographical location of the Study Area

In this experiment, we have taken PALSAR data of the Roorkee region of the Uttarakhand, India. The Roorkee is located at 29.87° N and 77.88° E. It has an elevation of 268 meters, i.e. 879 feet above the sea level, it is spread over the region of 129.88 sq. km [39], from the ArcGIS image we have obtained the initial information of the Roorkee region like water bodies in the form of Solani river, vegetation areas, urban settlements, areas where no vegetation is present and bare soil surfaces. Here we have classified the PALSAR data in the four classes, i.e. bare soil, urban, vegetation and water.
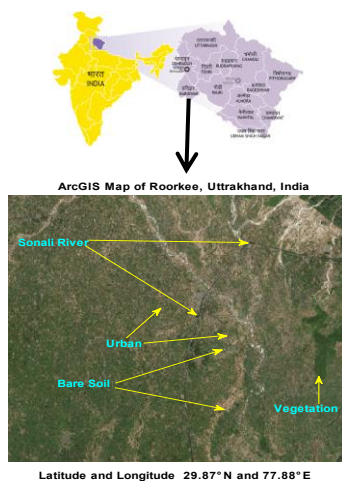

ArcGIS Map of Roorkee, Uttrakhand, India

Latitude and Longitude 29.87°N and 77.88°E

Fig.5 India (Bharat) → Uttarakhand→ Location of the Roorkee

Now for the training of the dataset we have manually selected region of interest (ROI's) for the four classes with the assistance of ArcGIS Earth, and for computing the accuracy for our classification we have used the ground truth coordinates obtained through field visits to the geographic locations. We have obtained the information of the particular class through latitude and longitude coordinates of the respective area with the assistance of GPS device.

S.No	Class	Reference ROI	Sample ROI	Allotted Color
1	Bare soil	100	20, 30, 40, 50,100, 150	⬛
2	Urban	100	20, 30, 40, 50, 100,150	🟥
3	Vegetation	100	30, 40, 50, 70,80,100,150	🟩
4	Water	100	20, 40,80,120, 160, 200	🟦

Table 1 Ground truth survey points

The PALSAR data is having importance over other datasets because it is quad-pole i.e. it provides information about four different polarizations HH (horizontal-horizontal), VV (vertical-vertical), HV (horizontal-vertical) and VH (vertical-horizontal), these polarimetric represents linear polarizations. Now we have four different bands of the PALSAR data characterize as Band 1, Band 2, Band 3 and Band 4, all these bands are fused together to obtain a quad polarized PALSAR image.

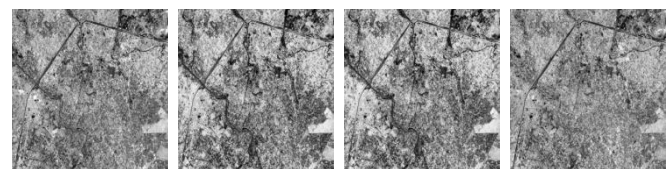(a) Band 1   (b) Band 2   (c) Band 3   (d) Band 4

Fig. 6 Different band representation of the PALSAR data

Now all the four bands are fused together to obtain a single quad pole image on which the classification schemes are applied. Map information of these bands is projection: geographic latitude/longitude, pixels: 0.00005 degree and Datum WGS-84, now all the bands are fused together to optically represent a PALSAR image shown in figure 7.

Fig.7 Band fused PALSAR image (Band 1+Band 2+ Bands 3+Band 4)

The purpose of this investigation is to identify the technique from which we can obtain maximum classification accuracy and relate this effect with the kappa coefficient. Here we have compared seven classification schemes on the same set of ROI's to obtain the maximum efficiency produced by an individual classification scheme. First of all, we have classified the PALSAR data through PC technique and the statistics derived after classification is shown in table 2 below.

#Notions used for Producer Accuracy= PA, User Accuracy=UA, Omission Error =OE, Commission Error=CE and Kappa Coefficient=KC

Classes	PA	UA	OE	CE
Vegetation	100.00	69.44	0.00	30.56
Urban	11.00	84.62	89.00	15.38
Water	47.00	83.93	53.00	16.07
Bare Soil	77.59	98.90	22.41	1.10

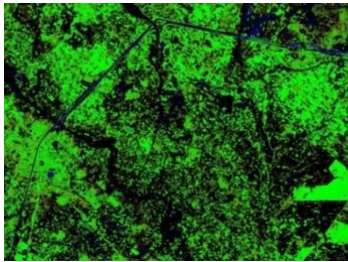Table 2 Statistics derived through Parallelepiped classification

Fig. 8 Parallelepiped classified PALSAR data

Here, after the classification we have obtained no useful information as the majority of the classified image is in green and black i.e. vegetation and bare soil respectively. The initial observation concludes that the region under water and urban classes remains unclassified. Thus this technique does not prove to be an effective technique for the image classification.

In the second step we have performed the classification of the PALSAR data through MDC technique. The statistics derived by performing this classification is shown in table 3.

Classes	PA	UA	OE	CE
Vegetation	73.00	90.68	27.00	9.32
Urban	86.00	60.14	14.00	39.86
Water	97.00	85.09	3.00	14.91
Bare Soil	81.90	96.94	18.10	3.06

Table 3 Statistics derived through Minimum distance classification
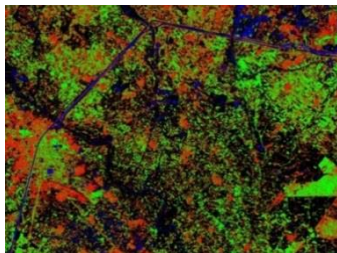

Fig. 9 Minimum distance classified PALSAR data

The MDC classification scheme proves to be a much better classification scheme as compared to the PC technique, as in MDC technique the initial assessment suggests that all four classes are classified water and urban class which were unclassified in the PC classification have also appeared in this classification scheme. In the third step, we have used MaDC technique for the classification of the PALSAR data. The statistics obtained after the classification of the data is shown in table 4.

Classes	PA	UA	OE	CE
Vegetation	83.00	82.18	17.00	17.82
Urban	63.00	64.08	34.00	35.92
Water	96.00	84.21	4.00	15.79
Bare Soil	79.31	94.85	20.69	5.15

Table 4 Statistics derived through Mahalanobis distance classification

The MaDC classification scheme appears to be an intermediate classification scheme in between PC and MDC classification scheme as from the classified image of MaDC we have come to a conclusion that all the four classes are classified, as all classes are making their presence count, but when comparing the classified image with MDC classified image. The urban class seems to be not completely classified.
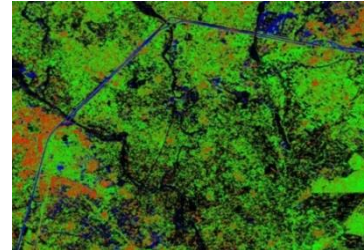

Fig. 10 Mahalanobis distance classified PALSAR data

In the fourth step, MLC scheme is used to classify the PALSAR data. The statistics derived after classification from MLC are shown in the table 5.

Classes	PA	UA	OE	CE
Vegetation	85.00	88.54	15.00	11.46
Urban	80.00	68.97	20.00	31.03
Water	78.00	87.64	22.00	12.36
Bare Soil	83.62	81.51	16.38	18.49

Table 5 Statistics derived through Maximum likelihood classification

MLC classification technique for image classification proves to be an effective classification technique as this technique provides us an better result in terms of visualization of the classification classes. This technique have certainly provided better classification result than PC and MaDC classification techniques. In this classification all the four classes are clearly visible, but the water stream and wetlands are not so clearly visible as compared to the classified image obtained from MDC classification.
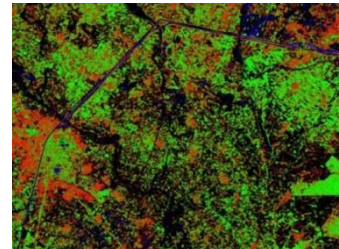

Fig. 11 Maximum likelihood classified PALSAR data

In the fifth classification step we have used SAM technique to classify the PALSAR image. The SAM classified PALSAR image is shown in the fig.12.
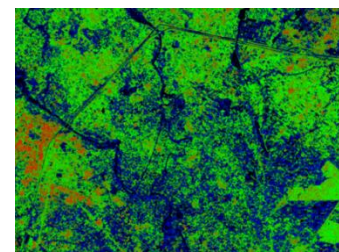

Fig. 12 Spectral angle mapper classified PALSAR data

Now after the classification of the image the statistics derived are represented in table 6.

Classes	PA	UA	OE	CE
Vegetation	86.00	72.27	14.00	27.73
Urban	62.00	68.13	38.00	31.87

Water	12.00	33.33	88.00	66.67
Bare Soil	51.72	95.24	48.28	4.76

Table.6 Statistics derived through Spectral angle mapper classification

SAM definitely proved to be a good classification technique but one thing is certainly true that it is not the best one. As in this classification scheme we have observed very less classified urban area, and at the same time, it is also observed that the class bare soil is also not so effectively classified. The image mostly appear to be bluish and greenish in colour which are the classes for water and vegetation respectively. This scheme is less accurate than the PC classification scheme. In the sixth classification scheme, we have used SVM classification scheme. This classification scheme does not provide effective classification information and the classified image is similar to PC classified image. The SVM classified image is shown in the fig.13.
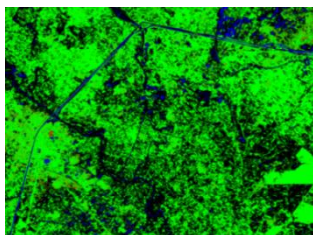

Fig. 13 Support vector machine classified PALSAR data

Now after the classification of the image the statistics derived are shown in table 7.

Classes	PA	UA	OE	CE
Vegetation	99.00	66.00	1.00	34.00
Urban	4.00	57.14	96.00	42.86
Water	98.00	83.05	2.00	16.95
Bare Soil	76.72	97.80	23.38	2.20

Table 7 Statistics derived through Support vector machine classification

SVM classification scheme is also less accurate classification scheme as from the classified image it is clear that all four classes are not completely classified. Here we have observed that urban region is not classified at all. The vegetation and bare soil classes are most significantly classified and class water is little bit classified. The urban class is completely unclassified. In the seventh and final classification scheme through SID, we have come to a conclusion that it is the worse classification scheme as classification results are tremendously weird. All the classes are completely unclassified and are not matched at all.
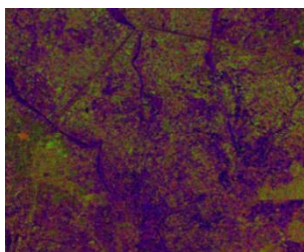

Fig. 14 Spectral information divergence classified PALSAR data

Now after the classification of the image the statistics derived are shown in table 8.

Classes	PA	UA	OE	CE
Vegetation	31.00	62.63	69.00	37.37
Urban	30.00	18.63	70.00	81.37
Water	15.00	15.63	85.00	84.38
Bare Soil	36.21	48.28	63.79	51.72

Table 8 Statistics derived through Spectral information divergence classification

SID classification scheme proves to be a worse classification scheme for the PALSAR data, as we have come to a conclusion that through this scheme no effective information can be obtained. The classified image only present a guess of the landcover in the mind of the user which appears only because we have used supervised classification techniques. This classification scheme is the least accurate and reliable in term of accuracy and kappa coefficient. Now, for obtaining the best classification accuracy we have compared the accuracy obtained from several classification techniques. The overall accuracy of the classification is obtained, which is the ratio of the correctly classified pixel to the total number of the pixels under investigation [40]. Here the value obtained through the kappa coefficient is also computed for the classification techniques, and on the basis of these both parameter, best classification technique is identified.

S.No	Classification Scheme	KC	Overall accuracy
1.	Parallelepiped Classification (PC)	0.5466	67.44 %
2.	Minimum distance classification (MDC)	0.7574	82.36%
3.	Mahalanobis distance Classification (MaDC)	0.7432	81.39 %
4.	Maximum likelihood classification (MLC)	0.7591	82.17%
5.	Spectral angle mapper (SAM)	0.4545	59.30%
6.	Support vector machine (SVM)	0.6423	75.38%
7.	Spectral information divergence (SID)	0.1011	28.87%

Table 9 Comparative analysis of kappa coefficient and overall accuracy

It is now clear from the table 9 that the range of the kappa coefficient is maximum for the MLC followed by MDC, MaDC, SVM, PC, SAM and minimum for SID. The overall accuracy is best for the MDC scheme, followed by MLC, MaDC, SVM, PC, SAM and minimum for SID classification scheme. Here we have also discovered a practical fact that the value of the kappa coefficient and overall efficiency is approximately equal for maximum likelihood and minimum distance classification. Now fig. 15 shows a comparative bar plot of the kappa values obtained from various classification techniques. The change obtained in the value of kappa coefficient for MLC and MDC is approx 0.2239 % which conclude that both the scheme can be used to produce the maximum kappa coefficent and overall accuracy.
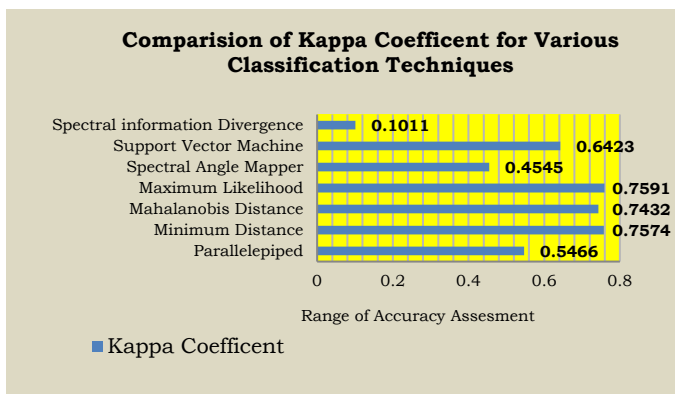
Fig. 15 Comparative barplot of kappa coefficient

In the fig. 16 comparative plot of the overall accuracy is shown. Finally, on comparing the obtained accuracy from various classification techniques, we have concluded that MDC technique have obtained maximum accuracy, which is 0.2352 % more than the MLC classification scheme.
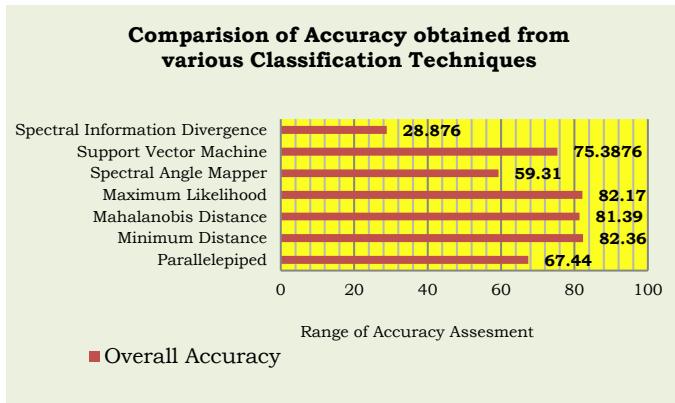

Fig. 16 Comparative barplot of the Overall Accuracy

Finally, we have developed a peculiar relation in between the obtained overall accuracy and obtained kappa coefficient for the various classification schemes, this relationship is linear in nature, from this we have also concluded that as the value of kappa coefficient increases the classification accuracy also get increased.

$$Kappa\ Coefficent \propto Overall\ Accuracy \qquad (24)$$

i.e. kappa coefficient is directly proportional to overall accuracy.

Now we have plotted the kappa values on vertical-axis and classification accuracy on the horizontal axis. This shows that a new linear relationship is developed in between kappa coefficient and overall classification accuracy.
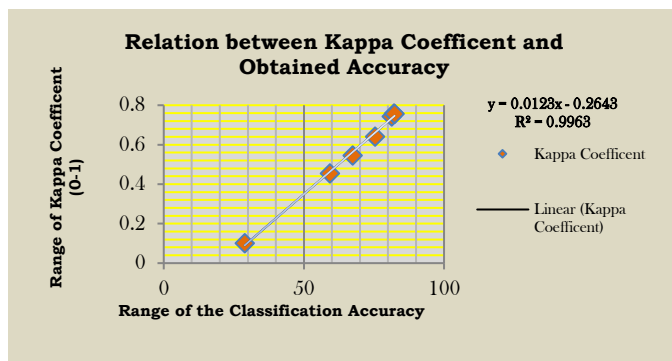

Fig. 17 Linear relationship between kappa coefficient and accuracy

5. CONCLUSION

The classification techniques proves to be a very effective and useful tool to match training and testing data, from this research work we have obtained conclusion that among the seven supervised classification techniques, MDC classification technique is best from the point of view of accuracy and MLC is best in term of kappa coefficient, furthermore MLC and MDC classification techniques both can provide good accuracy and kappa values. We have also concluded that classes like bare soil and vegetation are clearly visible through every classification scheme, but the classification of classes like urban and water is different from the different classification schemes. This can also be an area of research why water and urban are not able to appear clearly through various classification schemes. One important recommendation is that when any user wants his data to be classified to a minimum threshold then he can prefer SVM, PC, SAM over MLC, MDC and MaDC techniques as these techniques also provide fair enough classification of the PALSAR data.

ACKNOWLEDGEMENT

REFERENCES

1. Lu, D., Weng, Q.: A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing 28(5), 823–870 (March 2007)

2. Bagirov, A., Rubinov, A., Soukhoroukova, N. V., Yearwood, J.: Unsupervised and supervised data classification via nonsmooth and global optimization 11. Springer-Verlag, Australia (2003)

3. Rajendran, S., Al-Sayigh, A., Al-Awadhi, T.: Vegetation analysis study in and around Sultan Qaboos University, Oman, using Geoeye-1 satellite data. The Egyptian Journal of Remote Sensing and Space Science 19, 297-311 (April 2016)

4. Zhu, Z., Nandi, A.: Blind Digital Modulation Classification Using Minimum Distance Centroid Estimator and Non-Parametric Likelihood Function. IEEE Transactions on Wireless Communications 13(8), 4483 - 4494 (April 2014)

5. Mei, J., Liu, M., Wang, Y.-F., Gao, H.: Learning a Mahalanobis Distance based Dynamic Time Warping Measure for Multivariate Time Series Classification. IEEE Transactions on Cybernetics 46(6), 1363 - 1374 (June 2016)

6. Radhika, K., Varadarajan, S.: Satellite image classification of different

resolution images using cluster ensemble techniques. In : International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies, Chennai (2017)

7. Balas, C., Epitropou, G., Tsapras, A., Hadjinicolaou, N.: Hyperspectral imaging and spectral classification for pigment identification and mapping in paintings by El Greco and his workshop. Multimedia Tools and Applications 77(8), 9737-9751 (April 2018)

8. Zhang, E., Zhang, X., Yang, S., Wang, S.: Improving Hyperspectral Image Classification Using Spectral Information Divergence. Geoscience and Remote Sensing Letter 11(1), 249-253 (January 2014)

9. Garridoa, F., Verbekeb, W., Bravoc, C.: A Robust profit measure for binary classification model evaluation. Expert Systems with Applications 92, 154-160 (Feburary 2018)

10. Singha, A., Singh, K.: Satellite image classification using Genetic Algorithm trained radial basis function neural network, application to the detection of flooded areas. Journal of Visual Communication and Image Representation 42, 173-182 (January 2017)

11. Bu, D., Tuan, T., Hoang, N.-D., Thanh, N., Nguyen, D., Liem, N., Pradhan, B.: Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization. Landslides 14, 447-458 (April 2017)

12. Verma, A., Garg, P., Prasad, K. S.: Sugarcane crop identification from LISS IV data using ISODATA, MLC, and indices based decision tree approach. Arabian Journal of Geosciences 11(107) (January 2017)

13. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imageNet classification. In : International Conference on Computer Vision (ICCV), Santiago, Chile (2015)

14. Zhuang, H., Deng, K., Fan, H., Yu, M.: Strategies Combining Spectral Angle Mapper and Change Vector Analysis to Unsupervised Change Detection in Multispectral Images. IEEE Geoscience and Remote Sensing Letters 13(5), 681-685 (May 2016)

15. Hasan, E., Faginb, T., Alfy, Z., Hong, Y.: Spectral Angle Mapper and aeromagnetic data integration for gold-associated alteration zone mapping: a case study forthe Central Eastern Desert Egypt. International Journal of Remote Sensing 37(8), 1762-1776 (September 2015)

16. Renzaa, D., Martinez, E., Molina, I., M.Ballesteros, D.: Unsupervised change detection in a particular vegetation land cover type using spectral angle mapper. Advances in Space Research 59(8), 2019-2031 (April 2017)

17. Zhang, E., Zhang, X., Yang, S., Wang, S.: Improving Hyperspectral Image Classification Using Spectral Information Divergence. IEEE Geoscience and Remote Sensing Letter 11(1), 249- 253 (January 2014)

18. Khaleghi, M., Ranjbar, H., Shahabpour, J., Honarmand, M.: Spectral angle mapping, spectral information divergence, and principal component analysis of the ASTER SWIR data for exploration of porphyry copper mineralization in the Sarduiyeh area, Kerman province, Iran. Springer Applied Geomatics 6(1), 49-58 (March 2014)

19. Janati, M., Soulaimani, A., Admou, H., Youbi, N., Hafid, A., Hefferan, K.: Application of ASTER remote sensing data to geological mapping of basement domains in arid regions: a case study from the Central Anti-Atlas, Iguerda inlier, Morocco. Arabian Journal of Geosciences 7(6), 2407-2422 (June 2014)

20. Jiao, H., Zhong, Y., Zhang, L.: Artificial DNA Computing-Based Spectral Encoding and Matching Algorithm for Hyperspectral Remote Sensing Data. IEEE Transaction of Geo Science and Remote Sensing 50(10), 4085-4103 (October 2012)

21. Musci, M., Feitosa, R., Costa, G.: Assessment of Binary Coding Techniques for Texture Characterization in Remote Sensing Imagery.

IEEE Remote Sensing Letters 10(6), 1607-1611 (November 2013)

22. Sheng, G., Yang, W., Xu, T., Sun, H.: High-resolution satellite scene classification using a sparse coding based multiple feature combination. International Journal of Remote Sensing 33(8), 2395-2412 (Feburary 2012)

23. Shaoa, Y., S.Lunetta, R.: Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. ISPRS Journal of Photogrammetry and Remote Sensing 70, 78-87 (June 2012)

24. Bui, D., Tuan, T., Klempe, H., Pradhan, B., Revhaug, I.: Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides 13(2), 361-378 (April 2016)

25. Kia, M., Pirasteh, S., Pradhan, B., Mahmud, A., Sulaiman, W.: n artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. Download PDF 67(1), 251-264 (September 2012)

26. classification, A.: V.F. Rodriguez-Galiano; B. Ghimire; J. Rogan; M. Chica-Olmo; J.P. Rigol-Sanchez. ISPRS Journal of Photogrammetry and Remote Sensing 67, 93-104 (December 2011)

27. Maulik, U., Chakraborty, D.: Remote Sensing Image Classification: A survey of support-vector-machine-based advanced techniques. IEEE Geoscience and Remote Sensing Magzine 5(1), 33-52 (March 2017)

28. Hong, H., Liu, J., Zhu, A.-X., Shahabi, H., Pham, B., Chen, W., Pradhan, B.: A novel hybrid integration model using support vector machines and random subspace for weather-triggered landslide susceptibility assessment in the Wuning area (China). Envoirnmental Earth Science 19(201), 633-652 (October 2017)

29. Wang, X., Thomea, N., Cord, M.: Gaze latent support vector machine for image classification improved by weakly supervised region selection. Pattern Recognition 72, 59-71 (December 2017)

30. Schowengerdt, R.: Remote Sensing : Models and methods for image processing III edn. Elsevier U.K, U.K (2007)

31. Hasmadi, M., HZ, P., MF, S.: Evaluating supervised and unsupervised techniques for land cover mapping using remote sensing data. Malaysian Journal of Society and Space 5(1), 1-10 (2009)

32. Carletta, J.: Assessing agreement on classification tasks: the kappa statistic. ACM: Computational Linguistics 22(2), 249-254 (June 1996)

33. Shimada, M., Tadono, T., Rosenqvist, A.: Advanced land observing satellite (ALOS) and monitoring global environmental. Proceedings of the IEEE 98(5), 780-799 (May 2010)

34. Urai, M.: Time series analysis of discolored seawater reflectance observed by Advanced Visible and Near Infrared Radiometer type 2 (AVNIR-2) at Fukutoku-Okonaba submarine volcano, Japan. Journal of Volcanology and Geothermal Research 269, 23-27 (January 2014)

35. Takaku, J., Tadono, T.: Quality updates of 'AW3D' global DSM generated from ALOS PRISM. In : IEEE, Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, pp.5666-5669 (2017)

36. Sato, H., Suzuki, A.: Landslide Surface Deformation Detected by Synthetic Aperture Radar (SAR) Interferometry in Shizu Area on the Southern Foot of Mt. Gassan, Japan. GIS Landslide, 31-44 (May 2017)

37. Schlogel, R., Doubre, C., Malet, J.-P., Masson, F.: Landslide deformation monitoring with ALOS/PALSAR imagery: AD-InSAR geomorphological interpretation method. Geomorphology 231, 314–330 (2015)

38. Reichman, D., Collins, L., Malof, J.: On Choosing Training and Testing Data for Supervised Algorithms in Ground-Penetrating Radar Data for Buried Threat Detection. IEEE Transaction on Geo Science 56(1), 497 - 507 (October 2017)

39. Report, G.: District Census Handbook. Census of India , Directorate of Cencus Operations, Haridwar (2011)

40. Banko, G.: A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data and of Methods Including Remote Sensing Data in Forest Inventory. Interim , International Institute for Applied Systems Analysis , Laxenburg, Austria (1998)