Critical analysis of classification techniques for polarimetric synthetic aperture radar data

Vikas Mittal[a,1,*], Dharmendra Singh[b,2], Lalit Mohan Saini[c,3]

[a] *Electronics and Communication Engg. Dept., National Institute of Technology Kurukshetra, Kurukshetra, India*
[b] *Electronics and Communication Engg. Dept., Indian Institute of Technology Roorkee, Roorkee, India*
c Electrical Engg.Dept., National Institute of Technology Kurukshetra, Kurukshetra, India
[1] *vikas_mittal@nitkkr.ac.in* *; [2] *dharmfec1@gmail.com;* [3] *lmsaini@gmail.com*
* *corresponding author*

ARTICLE INFO

Article history:
Received March 26, 2016
Revised March 30, 2016
Accepted April 16, 2016

Keywords:
PolSAR features
Land cover classification
Supervised and unsupervised classification
Scattering mechanisms
Backscattering coefficients
Feature extraction and selection

ABSTRACT

Full polarimetry SAR data known as PolSAR contains information in terms of microwave energy backscattered through different scattering mechanisms (surface-, double- and volume-scattering) by the targets on the surface of land. These scattering mechanisms information is different in different features. Similarly, different classifiers have different capabilities as far as identification of the targets corresponding to these scattering mechanisms. Extraction of different features and the role of classifier are important for the purpose of identifying which feature is the most suitable with which classifier for land cover classification. Selection of suitable features and their combinations have always been an active area of research for the development of advanced classification algorithms. Fully polarimetric data has its own advantages because its different channels give special scattering feature for various land cover. Therefore, first hand statistics HH, HV and VV of PolSAR data along with their ratios and linear combinations should be investigated for exploring their importance vis-à-vis relevant classifier for land management at the global scale. It has been observed that individually first hand statistics yield low accuracies. And their ratios are also not improving the results either. However, improved accuracies are achieved when these natural features are stacked together

I. Introduction

The process of managing the use and development of land resources is known as land management. Land resources are the primitive resources which provide basis for further development activities. A variety of purposes such as agriculture, reforestation, water resource management and eco-tourism projects etc. use land resources [1].

Land management and study of patterns of consumption of other natural resources have become more relevant these days due to various reasons: climate change, growing population and cities, geographic and demographic distributions, costlier agricultural products etc. All these factors are increasing the pressure to make more intensive use of land area, which is gradually becoming short in supply, causing conflicts in various interests. One such conflict is “food versus fuel” simultaneously protecting the environment. The important question is that out of expansion of settlement areas and preservation of arable land which one should be given priority. These confusions are increasingly arising both at regional and global levels [2-7]. The solution of these problems and sustainable use of available land areas call for finding new techniques for land management. The new techniques should have wide scope and reach so that solutions at different scales can be evolved.

Radar polarimetry is one such technique which can provide dependable solutions to the above problems at matching scales. Interaction of radar frequencies with different land covers; categorized into various scattering mechanisms viz. single-, double- and volume-scattering provides an efficient and wide scope tool for land management. Radar can provide observations during night and inclement weather conditions i.e. presence of clouds, dust, rain etc. which is an advantage in comparison to other satellite sensors. Ease of availability of data from advanced synthetic aperture radar (SAR) sensors, such as RADARSAT, TerraSAR-X, ALOS PALSAR, RISAT etc. that operate in different range of microwave frequencies, led to further development of advanced techniques for land management using full polarimetric radar data (PolSAR) [8-10]. One of the basic steps for land management using radar polarimetry is land cover classification. Various land features such as water, urban, bare soil, vegetation etc. scatter incident radar signal through one of the above scattering mechanisms that manifest themselves differently for different polarizations HH, HV or VV. For instance, single-bounce is known to dominate in VV and double-bounce in HH polarization. Similarly their ratios, for example HV/VV or HV/HH, have been believed to be able to differentiate further these scattering mechanisms resulting in better identification of individual land cover [11].

Extraction and preprocessing of PolSAR data is followed by feature extraction suitable for an application. There are mainly three broad categories of PolSAR feature types; one based on original data and its transforms such as scattering-, coherent-matrix and backscattering intensity, second based on polarimetric target decompositions such as H/A/α, 3-, 4-component and Touzi decompositions and third based on other types such as texture and color etc as shown in Fig. 1 [12, 13]. Backscattering intensity is the primary and natural feature of PolSAR data obtained after its preprocessing and calibration without any approximations. It is converted into a physical quantity called backscattering coefficient expressed in dB. Typically its values range from +5 dBs for high to -40 dBs for low backscatter [14]. This feature is dependent on variety of factors related to both target and illuminating radar such as size, shape, orientation and type of scatterers, moisture contents in the illuminated targets and frequency, polarization and incidence angle of radar signal. All other features are derived or transformed features based on certain mathematical assumptions.

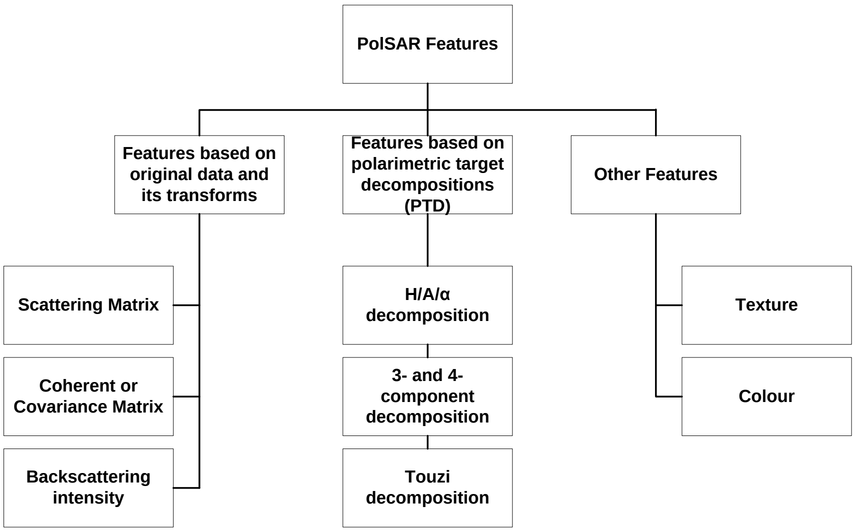

Fig. 1 Broad classification of PolSAR features

During early years of PolSAR i.e. second last decade of twentieth century, land cover classification studies were conducted using intensity features but with low classification accuracies [12, 15]. Later on, it was tried to increase classification accuracies using other transformed and decomposed features. In general, they have been used in conjunction with other features such as texture [16] etc. on one hand and myriad combination of classification and labeling algorithms to improve classification accuracies. This made the whole process of radar polarimetry based land cover classification complicated with requirement of additional inputs in terms of prior information, which is not always readily available. Availability or absence of prior information will decide the classification approach to be used; supervised in the first case and unsupervised in the latter case. Also the process can be simplified by relying on first order statistics of land cover classification. In any case, development of autonomous classification techniques with minimum human intervention is desirable. Hence, in the present study backscattering coefficients and their combinations are investigated comprehensively for their suitability for land cover classification using various supervised and unsupervised classifiers. This will help in assessing which of these are able to better identify which land cover on its own without any additional information.

II. Study Area and Data Used

The study region is lying between $29^{\circ}49'47.45'' - 29^{\circ}54'0.19''N$ to $77^{\circ}50'59.29'' - 77^{\circ}55'12.03''E$. The study region mainly consists of Roorkee city situated in the state of Uttarakhand, India. The city contains urban area, a rivulet Solani in the north. Ganga canal divides the whole city into two equal halves. The study area has abundant of land cover such as water, urban, wetland, bare soil, short vegetation and tall vegetation. Ganga canal is the main water body and regions around the banks of rivulet Solani constitutes wetland cover. Baresoil exists in the open areas in and around city. Short and long vegetation spread over the whole study region.

ALOS PALSAR fully polarimetric data observed on 9th April 2010is used here. This product is L1.1 data in CEOS format with scene ID ALPSRP224150590. It is a single look complex data on slant range and has seven numbers of looks in azimuth. The terrain of the study area is flat (slope<1°), hence backscattering coefficient is calculated using a constant incident angle (23.989°) due to small deviation in it (22.5° – 24°) [17].

Over the whole study region, ground sample points (GSPs) of each land cover i.e. water, urban, wetland, bare soil, short vegetation and tall vegetation are selected. Various classification accuracies are computed using testing ground samples. The ground sample points are selected through extensive ground survey and Google Earth.

III. Proposed Methodology

A. Data Preprocessing

Pre-processing of full polarimetric ALOS PALSAR data is done to extract different datasets for *HH*, *HV* and *VV* polarizations. These datasets contain normalized backscattering coefficients which are randomly distributed over the whole real numbers. ENVI software and its SARSCAPE module are used for pre-processing PALSAR data using steps outlined by Mishra et al. [10]. Focused PALSAR data is directly imported to extract single look complex (SLC) files which are multi-looked by a factor of 7 to improve radiometric resolution due to different resolutions in range and azimuth directions. Digital elevation model (DEM) is extracted using GO TOPO30 for the purpose of geo-referencing by nearest neighbor approximation. Phase-mod form of the geo-referenced data is then converted to complex files which are separated into real and imaginary parts. Band Math is then applied to extract normalized backscattering coefficient using the formula given in [10]. The process is repeated for each polarization.

B. PolSAR features

After preprocessing of PolSAR data, various datasets *HH*, *HV* and *VV* are obtained. These datasets contain backscattered intensity values calibrated in terms of backscattering coefficients σ_{hh}, σ_{hv} and σ_{vv}, respectively. These are expressed in dBs. Different scattering mechanisms i.e. surface- (bare soil, wetland, water etc.), double bounce- (urban, buildings etc.), volume-scattering (tall vegetation, forest etc.) present themselves prominently in different features. For instance, double bounce-, surface- and volume-scattering is more prominent in HH, VV and HV features, respectively.

The other features are obtained by taking their ratios $\frac{HH}{HV}, \frac{HH}{VV}, \frac{HV}{VV}, \frac{HV}{HH}, \frac{VV}{HH}$ and $\frac{VV}{HV}$. Different researchers have found that a particular land cover class is highlighted by these ratio features [11]. Linear combinations of HH, HV and VV (e.g. HH±HV etc.) are also considered for their effect on identification of various land covers. But still role of classifier is to be investigated regarding which feature is suitable for identifying which land cover.

All the above features correspond to single bands or their ratios. Some of supervised classifiers such as Mahanalobis distance (MD) and maximum likelihood classifier (MLE) requires at-least two bands. Therefore some of the features are stacked to use them. For instance, σ_{hh} is chosen to be stacked with $\frac{\sigma_{hv}}{\sigma_{vv}}$ or $\frac{\sigma_{vv}}{\sigma_{hv}}$ as it is not appearing in either ratio. Similarly, stacking of all three σ_{hh}, σ_{hv} and σ_{vv} is investigated (Table 1). Physical significance of some of these features is summarized in [11] (the list is not exhaustive due to still unknown nature of these complex phenomena).

Table 1. Physical significance of backscattering coefficients [11]

Feature	Sensitivity to scattering mechanism	Expected land cover identification
σ_{hh}	double-bounce	Buildings and high structures
σ_{hv}	multiple-bounce	Vegetation, bio-mass etc.
σ_{vv}	single-bounce	Water, smooth bare soil etc.
$\frac{\sigma_{hv}}{\sigma_{vv}}$ and $\frac{\sigma_{vv}}{\sigma_{hv}}$	Differentiates single- and multiple-bounce	Discriminates bare soil and vegetation

C. *Supervised Classification*

Supervised classification requires representative ground sample points corresponding to different land covers. These sample land cover classes are called “training sites”. Based on these training sites the supervised classifier identifies various land cover classes in the entire dataset. It assigns each data point to one of the land covers it resembles most in the training set. The common supervised classifiers are parallelepiped, minimum distance, Mahanalobis distance and maximum likelihood.

Parallelepiped classification uses decision boundaries forming an n-dimensional parallelepiped. Its dimensions are defined using a standard deviation threshold from the mean of each selected class. If a pixel value lies between the lower and upper thresholds, it is assigned to that class. Areas that do not fall within any of the parallelepiped classes are designated as unclassified.

Minimum distance classifier calculates the Euclidean distance from each unknown pixel to the mean vector for each class. All pixels are classified to the nearest (minimum distance) class. The Mahanalobis distance classification is a direction-sensitive distance classifier that uses statistics for each class. It overcomes some of the limitations of minimum distance classifier that arise in case of poorly scaled and highly correlated features. This classifier uses Mahanalobis distance measure instead of Euclidean distance which takes into account number of standard deviations that a point is away from each class mean along different principal component axis instead of simple distance. Therefore, this measure is without units and scale-invariant. Corresponding to rescaling to unit variance along each axis, Mahanalobis distance is identical to Euclidean distance.

Mahanalobis distance classifier is similar to maximum likelihood classifier for equal class covariances. On the other hand, maximum likelihood classifier works on the assumption of normal distribution of each class statistics. It calculates the probability that a given pixel belongs to a specific class. Each data point is assigned to the class that has the highest probability (maximum likelihood). All data points are classified for no probability threshold [18].

Class-wise producer’s and user’s classification accuracies for some of the features for different supervised classifiers are summarized in Table 2 to Table 7.

The accuracies in Table 2 to Table 7 given as NR are either zero, negative or abysmally low, hence these values does not have any practical significance. Stacking of HH with ratio of cross features i.e. HV/VV or VV/HV does not improve classification accuracy.

Since Mahanalobis and Maximum likelihood classifier require at-least two bands, therefore different features are stacked for land cover classification using them. Various classification accuracies for one of the features obtained by stacking HH, HV, VV for different supervised classifiers are shown in Table 2 to Table 7. Entries corresponding to single features are marked NF in the above tables.

Table 2. Classification Accuracies of 'water'with Supervised Classifiers

Classifier / Feature	Parallelepiped		Minimum Distance		Mahanalobis Distance		Maximum Likelihood	
	PA	***UA***	***PA***	***UA***	***PA***	***UA***	***PA***	***UA***
HH	68.18	31.38	67.27	49.66	NF	NF	NF	NF
HV	68.18	26.32	56.36	41.33	NF	NF	NF	NF
VV	67.27	37.19	67.27	55.64	NF	NF	NF	NF
HH/HV	68.18	16.52	4.55	12.82	NF	NF	NF	NF
HH/VV	NR	NR	4.55	12.82	NF	NF	NF	NF
HV/HH	68.18	16.52	21.82	16.44	NF	NF	NF	NF
HV/VV	68.18	16.52	33.64	21.51	NF	NF	NF	NF
VV/HH	NR	NR	4.55	12.82	NF	NF	NF	NF
VV/HV	73.64	19.10	33.64	21.51	NF	NF	NF	NF
HH+HV	68.18	36.23	66.36	58.40	NF	NF	NF	NF
Stacked(HH,HV,VV)	33.64	46.25	68.18	65.79	65.45	62.07	61.82	67.33

PA & UA – Producer's and user's accuracy (%), NR – not relevant, NF – not feasible

Table 3. Classification Accuracies of 'urban' with Supervised Classifiers

Classifier / Feature	Parallelepiped		Minimum Distance		Mahanalobis Distance		Maximum Likelihood	
	PA	***UA***	***PA***	***UA***	***PA***	***UA***	***PA***	***UA***
HH	64.50	36.13	64.50	57.59	NF	NF	NF	NF
HV	69	64.49	71.50	84.62	NF	NF	NF	NF
VV	71.50	36.11	6.50	24.53	NF	NF	NF	NF
HH/HV	23.50	34.31	2.50	26.32	NF	NF	NF	NF
HH/VV	NR	NR	2.50	26.32	NF	NF	NF	NF
HV/HH	23.50	34.31	12	37.50	NF	NF	NF	NF
HV/VV	23.50	34.31	54.50	48.02	NF	NF	NF	NF
VV/HH	NR	NR	2.50	26.32	NF	NF	NF	NF
VV/HV	32.50	44.22	54.50	48.02	NF	NF	NF	NF
HH+HV	67	52.76	70.50	75.40	NF	NF	NF	NF
Stacked(HH,HV,VV)	32.50	73.03	72.5	84.3	63	82.89	70	90.91

PA & UA – Producer's and user's accuracy (%), NR – not relevant, NF – not feasible

Table 4. Classification Accuracies of 'wetland' with Supervised Classifiers

Classifier / Feature	Parallelepiped		Minimum Distance		Mahanalobis Distance		Maximum Likelihood	
	PA	***UA***	***PA***	***UA***	***PA***	***UA***	***PA***	***UA***
HH	7.64	12.36	25	36	NF	NF	NF	NF
HV	NR	NR	25	43.37	NF	NF	NF	NF
VV	NR	NR	47.22	33.66	NF	NF	NF	NF
HH/HV	46.53	65.05	52.78	25.85	NF	NF	NF	NF
HH/VV	NR	NR	52.78	25.85	NF	NF	NF	NF
HV/HH	46.53	65.05	77.08	55.22	NF	NF	NF	NF
HV/VV	46.53	65.05	84.72	75.78	NF	NF	NF	NF
VV/HH	NR	NR	52.78	25.85	NF	NF	NF	NF
VV/HV	68.75	77.95	84.72	75.78	NF	NF	NF	NF
HH+HV	56.25	31.76	15.97	29.11	NF	NF	NF	NF
Stacked(HH,HV,VV)	33.33	94.12	86.11	86.05	87.5	75.9	88.89	79.01

PA & UA – Producer's and user's accuracy (%), NR – not relevant, NF – not feasible

Vol. 2, No. 1, March 2016, pp. 7-17

Table 5. Classification Accuracies of 'bare soil' with Supervised Classifiers

Classifier / Feature	Parallelepiped		Minimum Distance		Mahanalobis Distance		Maximum Likelihood	
	PA	***UA***	***PA***	***UA***	***PA***	***UA***	***PA***	***UA***
HH	3.16	10.34	20	15.45	NF	NF	NF	NF
HV	65.26	27.56	42.11	31.25	NF	NF	NF	NF
VV	25.26	21.82	21.05	17.09	NF	NF	NF	NF
HH/HV	NR	NR	50.53	16.16	NF	NF	NF	NF
HH/VV	NR	NR	50.53	16.16	NF	NF	NF	NF
HV/HH	NR	NR	12.63	14.12	NF	NF	NF	NF
HV/VV	NR	NR	12.63	14.63	NF	NF	NF	NF
VV/HH	NR	NR	50.53	16.16	NF	NF	NF	NF
VV/HV	NR	NR	12.63	14.63	NF	NF	NF	NF
HH+HV	NR	NR	22.11	21.65	NF	NF	NF	NF
Stacked(HH,HV,VV)	32.63	30.10	41.05	35.78	40	36.54	47.37	38.79

PA & UA – Producer's and user's accuracy (%), NR – not relevant, NF – not feasible

Table 6. Classification Accuracies of 'short vegetation' with Supervised Classifiers

Classifier / Feature	Parallelepiped		Minimum Distance		Mahanalobis Distance		Maximum Likelihood	
	PA	***UA***	***PA***	***UA***	***PA***	***UA***	***PA***	***UA***
HH	NR	NR	26.67	36.36	NF	NF	NF	NF
HV	NR	NR	36	41.22	NF	NF	NF	NF
VV	NR	NR	28.67	44.33	NF	NF	NF	NF
HH/HV	NR	NR	12	27.27	NF	NF	NF	NF
HH/VV	NR	NR	12	27.27	NF	NF	NF	NF
HV/HH	NR	NR	11.33	32.69	NF	NF	NF	NF
HV/VV	NR	NR	14	27.27	NF	NF	NF	NF
VV/HH	NR	NR	12	27.27	NF	NF	NF	NF
VV/HV	NR	NR	14	27.27	NF	NF	NF	NF
HH+HV	NR	NR	38	37.75	NF	NF	NF	NF
Stacked(HH,HV,VV)	18	61.36	40.67	52.59	32.67	52.69	40	51.72

PA & UA – Producer's and user's accuracy (%), NR – not relevant, NF – not feasible

Table 7. Classification Accuracies of 'tall vegetation' with Supervised Classifiers

Classifier / Feature	Parallelepiped		Minimum Distance		Mahanalobis Distance		Maximum Likelihood	
	PA	***UA***	***PA***	***UA***	***PA***	***UA***	***PA***	***UA***
HH	NR	NR	14.67	16.18	NF	NF	NF	NF
HV	NR	NR	52	34.51	NF	NF	NF	NF
VV	NR	NR	25.33	11.05	NF	NF	NF	NF
HH/HV	1.33	25	8	10.17	NF	NF	NF	NF
HH/VV	NR	NR	8	10.17	NF	NF	NF	NF
HV/HH	1.33	25	53.33	17.70	NF	NF	NF	NF
HV/VV	1.33	25	4	5.45	NF	NF	NF	NF
VV/HH	NR	NR	8	10.17	NF	NF	NF	NF
VV/HV	NR	NR	4	5.45	NF	NF	NF	NF
HH+HV	NR	NR	36	20	NF	NF	NF	NF
Stacked(HH,HV,VV)	17.33	34.21	49.33	33.64	49.33	25.87	62.67	37.60

PA & UA – Producer's and user's accuracy (%), NR – not relevant, NF – not feasible

D. Unsupervised Classification

Unsupervised classification does not require any or minimum a-prior information for identification of land cover classes. K-means and ISODATA are most commonly used classifiers for

PolSAR land cover classification in an unsupervised manner. Both are basically iterative clustering algorithms.

In general, an arbitrary initial cluster is assigned in both the algorithms. Then each pixel is assigned to the closest cluster. After that new cluster means are calculated for all the pixels in a cluster. The above two steps are repeated until there is a little "change" between the iterations.

The ISODATA algorithm is more refined than K-means in terms of splitting and merging of clusters. Merging is allowed if numbers of pixels in a cluster are less than or if the centers of two clusters are closer than a certain threshold. Similarly splitting is allowed for the case of standard deviation more than a predefined threshold or numbers of pixels are twice than the agreed threshold [18].

The ISODATA algorithm is unlike K-means algorithm in the sense that in the former different number of clusters are allowed while latter assumes number of clusters a-priori.

Class-wise producer's and user's classification accuracies for some of the features for different unsupervised classifiers are summarized in Table 8 to Table 13.

Table 8. Classification Accuracies of 'water' with unsupervised Classifiers

Classifier / **Feature**	**K-means**		**ISODATA**	
	PA	***UA***	***PA***	***UA***
HH	38.18	57.53	38.18	57.53
HV	36.36	54.05	36.36	54.05
VV	51.82	65.52	51.82	65.52
HH+HV+VV	44.45	72.06	44.55	72.06
Stacked (HH,HV,VV)	54.55	75	54.55	75

Table 9. Classification Accuracies of 'urban' with unsupervised Classifiers

Classifier / **Feature**	**K-means**		**ISODATA**	
	PA	***UA***	***PA***	***UA***
HH	26	71.23	26.00	71.23
HV	67.5	87.1	67.5	87.1
VV	34	56.20	24	56.2
HH+HV+VV	41.5	77.57	41.5	77.57
Stacked (HH,HV,VV)	47.5	87.16	47.5	87.16

Table 10. Classification Accuracies of 'wetland' with unsupervised Classifiers

Classifier / **Feature**	**K-means**		**ISODATA**	
	PA	***UA***	***PA***	***UA***
HH	44.44	32.82	44.44	32.82
HV	40.97	42.75	40.97	42.75
VV	57.64	38.07	57.64	38.07
HH+HV+VV	43.06	28.97	39.58	33.53
Stacked (HH,HV,VV)	66.67	51.89	66.67	51.89

Table 11. Classification Accuracies of 'bare soil' with unsupervised Classifiers

Classifier / **Feature**	**K-means**		**ISODATA**	
	PA	***UA***	***PA***	***UA***
HH	37.89	18.65	37.89	18.65
HV	43.16	28.87	43.16	28.87
VV	23.16	14.97	23.16	14.97
HH+HV+VV	33.68	25.2	33.68	25.20
Stacked (HH,HV,VV)	42.11	26.14	4.21	4.71

Table 12. Classification Accuracies of ‘short vegetation’ with unsupervised Classifiers

Classifier / Feature	K-means		ISODATA	
	PA	***UA***	***PA***	***UA***
HH	32.67	34.03	32.67	34.03
HV	32.67	37.12	32.67	37.12
VV	26	41.49	26	41.49
HH+HV+VV	26.67	45.45	26.67	45.45
Stacked (HH,HV,VV)	13.33	23.53	47.33	46.41

Table 13. Classification Accuracies of ‘tall vegetation’ with unsupervised Classifiers

Classifier / Feature	K-means		ISODATA	
	PA	***UA***	***PA***	***UA***
HH	8	6.25	8	6.25
HV	58.67	33.08	58.67	33.08
VV	18.67	13.08	18.67	13.08
HH+HV+VV	25.33	11.18	50.67	17.76
Stacked (HH, HV,VV)	66.67	30.86	66.67	30.86

After rigorous iterations, it was found that the following parameters for K-means and ISODATA unsupervised classifiers give the best classification results; no. of iterations=50, change threshold=4, classes=6 (both K-means and ISODATA), no. of pixels=10 (for ISODATA). No maximum standard deviation from mean and maximum distance error has been specified in K-means while maximum class deviation and minimum class distance are taken to be 1 and 5, respectively for ISODATA classifier.

E. Critical Analysis and Discussion

Various natural PolSAR features, their ratios and linear combinations have been investigated for various land covers (water, urban, wetland, bare soil, short- and tall-vegetation) classification using various supervised (parallelepiped, minimum distance, Mahanalobis distance and maximum likelihood classifier) and unsupervised (ISODATA and K-means) classifiers. Selection of features and their combinations has always been an active subject of investigation in PolSAR land cover classification studies. Various advanced computational algorithms have been applied for proper selection of the desired features [11].

However, in this study first hand statistics of PolSAR and their linear combinations are considered for land cover classification. This will help in assessing suitability of different features for this application along with the role of classifier. This is expected to lead to the end objective of developing generic unsupervised algorithms for the application of land cover classification with minimum human intervention. For comparison various supervised classifiers have also been taken into account.

The class ‘water’ is identified with a PA of around 68% using all features mentioned in Table 2 with the exception of stacked (HH,HV,VV), HH/VV and VV/HH features using parallelepiped classifier. The accuracies are abysmally low for HH/VV and VV/HH features. But UA is low for all the features. The minimum distance classifier results in PA of around 68% for HH, VV, HH+HV and stacked (HH, HV, VV) features with reasonable high UA for last two features. Since Mahanalobis distance and maximum likelihood classifiers work with at-least two bands so entries corresponding to single features are shown as not-feasible (NF) in Table 2. PA and UA are obtained in the range of 62-65% and 62-67% respectively for these two classifiers. In case of unsupervised classification, highest PA and UA of 55% and 75% are obtained for stacked (HH, HV, VV) feature for both K-means and ISODATA classifiers as shown in Table 8. Therefore, it is observed that stacked (HH, HV, VV) feature with minimum distance supervised classifier gives overall good results for water land cover. And, both unsupervised classifiers give identical results for this class.

Urban land cover is classified with the highest PA of 71.5% for VV and highest UA of 73.03% for stacked (HH, HV and VV) with parallelepiped classifier as shown in Table 3. The same trend is observed for minimum distance classifier with highest PA of 72.5% and UA of 84% for stacked (HH, HV, VV). However, in this case UA is highest for HV feature. Maximum likelihood classifier results in very good UA of about 91% for this class. In case of unsupervised classification, both K-means and ISODATA result in highest PA of 67.5% and UA of 87.1% for HV feature as observed from Table 9. In nutshell, it is observed that stacked (HH, HV and VV) feature is good for identifying urban land cover either with minimum distance or maximum likelihood classifier. On the other hand, HV feature can be used with either K-means or ISODATA classifier for urban land cover.

Wetland land cover is classified with highest PA of 68.75% and UA of 77.95% for VV/HV with parallelepiped classifier as observed from Table 4. However, UA of 94.12 % is highest for stacked (HH, HV and VV) feature which also give highest PA of 86.11% and UA of 86.05% with minimum distance classifier. PA of 88.89% is highest for stacked (HH, HV and VV) feature with maximum likelihood classifier. PA of 66.67% and UA of 51.89% are highest for stacked (HH, HV and VV) feature with both K-means and ISODATA unsupervised classifier as observed from Table 10. Therefore, stacked (HH, HV and VV) feature works best with minimum distance or maximum likelihood supervised classifiers and both K-means and ISODATA unsupervised classifiers for identifying wetland land cover.

HV feature gives best PA of 65.26% with parallelepiped classifier while it is around 50% for HH/HV, HH/VV and VV/HH features with minimum distance classifier as observed from Table 5. UA of 27.56% is obtained in the former case and UA of 16.16% is obtained in the latter case. Both these values are quite low. Similar trend is observed for other two supervised classifiers i.e. Mahanalobis and maximum likelihood classifier. HV also give highest PA of 43.16% and UA of 28.87% with both K-means and ISODATA unsupervised classifiers. Therefore, it can be summarized that HV feature can be used for this class with parallelepiped classifier, K-means or ISODATA classifier.

Very poor classification results are obtained for short vegetation class with parallelepiped classifier as is evident from Table 6. However, minimum distance classifier results in PA of 40.67% and UA of 52.59% for stacked (HH, HV and VV) feature. Same PA but different UA of 51.72% is obtained for this feature with maximum likelihood classifier. ISODATA unsupervised classifier gives best PA of 47.33% and UA of 46.41% for stacked (HH, HV and VV) feature as is observed from Table 12. Therefore, short vegetation can be identified best with stacked (HH, HV and VV) feature using either minimum distance classifier in supervised approach or ISODATA classifier in unsupervised approach.

Performance of parallelepiped classifier is also not good for any of the features for tall vegetation land cover as is observed from Table 7. The minimum distance classifier gives best PA of 53.33% for HV/HH feature but with low UA of only 17.7%. Both PA and UA are reasonably high for stacked (HH, HV and VV) feature. However, PA of 62.67% is highest for maximum likelihood classifier. Stacked (HH, HV and VV) feature also results in highest PA of 66.67% for both K-means and ISODATA classifiers as observed from Table 13. Therefore, stacked (HH, HV and VV) feature seems to be appropriate with either maximum likelihood supervised classifier or K-means or ISODATA unsupervised classifiers.

It has been observed from above discussion that not any single feature or classifier is appropriate for identifying different land covers. In general, the ratios and linear combinations of nature PolSAR features are not giving encouraging results. On the other hand, when they are stacked as in stacked (HH, HV and VV) feature, the good classification results are obtained for most of the land covers using both supervised and unsupervised approaches. Not very high values of classification accuracies can be achieved because in this study no additional information other than natural PolSAR features are taken into account.

The bigger challenge, however, is land cover classification in the scenario when minimum or no prior information is available. Only unsupervised classification can be used in such cases. Therefore, to make the study complete a comparison with other popular unsupervised classifier viz. Wishart classifier based on H/A/α decomposed features is also shown in Table 14.

Table 14. Classification Accuracies with Wishart unsupervised classifier

Class / Feature	Water		Urban		Wetland		Baresoil		Short Veg		Tall Veg	
	PA	***UA***	***PA***	***UA***	***PA***	***UA***	***PA***	***UA***	***PA***	***UA***	***PA***	***UA***
H/A/α	51.67	75.61	66.67	100	77.42	60.76	53.73	55.38	59.21	55.56	NR	NR

NR – not relevant

Classification accuracies shown in Table 14 are obtained without any additional information. These accuracies are better than stacked (HH, HV and VV) feature used in this study. This can be attributed to the fact that Wishart unsupervised classifier exploits the Wishart statistical distribution of PolSAR data which is not present in stacking of features.

IV. Conclusion

In this study, various PolSAR features used for land cover classification have been summarized. Natural features, their ratios and linear combinations have been investigated using various supervised and unsupervised classification for their suitability for land cover classification using PolSAR data. It is concluded from this study that the natural features HH, HV and VV stacked together may give improved classification results in both supervised and unsupervised domain. Further explorations are required to be done for enhancing the information content of PolSAR data and select/design suitable classifier for improved land cover classification keeping the two ends i.e. features and classifiers simple computationally and algorithmically.

References

[1] P. K. Mallupattu and J. R. Sreenivasula Reddy, "Analysis of Land Use/Land Cover Changes Using Remote Sensing Data and GIS at an Urban Area, Tirupati, India," The Scientific World Journal, vol. 2013, p. 6, 2013.

[2] J. Chilar, "Land cover mapping of large areas from satellites: Status and research priorities," International Journal of Remote Sensing, vol. 21, pp. 1093-1114, 2000/01/01 2000.

[3] K. B. Jones, "Importance of Land Cover and Biophysical Data in Landscape-Based Environmental Assessments," in North America Land Cover Summit, J. C. Campbell, K. B. Jones, J. H. Smith, and M. T. Koeppe, Eds., ed. National Academy of Sciences, Washington D.C.: Association of American Geographers, 2006, pp. 215-250.

[4] V. Alberga, "A study of land cover classification using polarimetric SAR parameters," International Journal of Remote Sensing, vol. 28, pp. 3851-3870, 2007/09/10 2007.

[5] H. T. Chu and L. Ge, "Improvement of land cover classification performance in western Australia using multisource remote sensing data," in 7th International Symposium on Digital Earth (ISDE7), Perth, Australia, 2011.

[6] S. K. Deb and R. K. Nathr, "Land use/cover classification - An introduction review and comparison," Global Journal of Researches in Engineering Civil and Structural Engineering, vol. 12, 2012.

[7] C. Qiang, K. Gangyao, J. Li, S. Lichun, and L. Diangang, "Unsupervised Land Cover/Land Use Classification Using PolSAR Imagery Based on Scattering Similarity," Geoscience and Remote Sensing, IEEE Transactions on, vol. 51, pp. 1817-1825, 2013.

[8] P. Ferrazzoli, S. Paloscia, P. Pampaloni, G. Schiavon, S. Sigismondi, and D. Solimini, "The potential of multifrequency polarimetric SAR in assessing agricultural and arboreous biomass," IEEE Transactions on Geoscience and Remote Sensing, vol. 35, pp. 5-17, 1997.

[9] A. Rodrigues, D. G. Corr, E. Pottier, L. Ferro-Famil, and D. Hoekman, "Land Cover Classification using Polarimetric SAR data," presented at the Workshop on applications of SAR polarimetry and polarimetric interferometry (POLInSAR 2003), ESA-ESRIN Frascati Italy, 2003.

[10] P. Mishra, D. Singh, and Y. Yamaguchi, "Land Cover Classification of Palsar Images by Knowledge Based Decision Tree Classifier and Supervised Classifiers Based on SAR Observables," PIER B, 2011.

[11] P. Mishra and D. Singh, "A Statistical-Measure-Based Adaptive Land Cover Classification Algorithm by Efficient Utilization of Polarimetric SAR Observables," Geoscience and Remote Sensing, IEEE Transactions on, vol. 52, pp. 2889-2900, 2014.

[12] V. Alberga, G. Satalino, and D. K. Staykova, "Comparison of polarimetric SAR observables in terms of classification performance," International Journal of Remote Sensing, vol. 29, pp. 4129-4150, 2008/07/01 2008.

[13] V. Alberga, D. Staykova, E. Krogager, A. Danklmayer, and M. Chandra, "Comparison of methods for extracting and utilizing radar target characteristic parameters," in Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Proceedings. 2005 IEEE International, 2005, pp. 2019-2021.

[14] P. S. Thenkabail, Land Resources Monitoring, Modeling, and Mapping with Remote Sensing: CRC Press, 2015.

[15] I. G. Cumnling and J. J. v. Zyl, "Feature Utility In Polarimetric Radar Image Classificatiion," in Geoscience and Remote Sensing Symposium, 1989. IGARSS'89. 12th Canadian Symposium on Remote Sensing., 1989 International, 1989, pp. 1841-1846.

[16] E. Nezry and F. Yakam-Simen, "On the preservation of polarimetric signatures and polarimetric texture signatures by fully polarimetric MAP filters," in Geoscience and Remote Sensing Symposium, 1999. IGARSS '99 Proceedings. IEEE 1999 International, 1999, pp. 1555-1557 vol.3.

[17] G. Mittal, P. Mishra, and D. Singh, "Fusion of polarimetric channel information of PALSAR data for land cover classification," Geomatics, Natural Hazards and Risk, vol. 5, pp. 157-172, 2014/04/03 2014.

[18] J. A. Richards and X. Jia, Remote Sensing Digital Image Analysis: An Introduction: Springer-Verlag New York, Inc., 1999.