Progress In Electromagnetics Research B, Vol. 30, 47–70, 2011

LAND COVER CLASSIFICATION OF PALSAR IMAGES BY KNOWLEDGE BASED DECISION TREE CLASSIFIER AND SUPERVISED CLASSIFIERS BASED ON SAR OBSERVABLES

P. Mishra and D. Singh

Department of Electronics and Computer Engineering
Indian Institute of Technology Roorkee
Roorkee, (UK)-247667, India

Y. Yamaguchi

Department of Information Engineering
Niigata University, Japan

Abstract—The intent of this paper is to explore the application of information obtained from fully polarimetric data for land cover classification. Various land cover classification techniques are available in the literature, but still uncertainty exists in labeling various clusters to their own classes without using any a priori information. Therefore, the present work is focused on analyzing useful intrinsic information extracted from SAR observables obtained by various decomposition techniques. The eigenvalue decomposition and Pauli decomposition have been carried out to separate classes on the basis of their scattering mechanisms. The various classification techniques (supervised: minimum distance, maximum likelihood, parallelepiped and unsupervised: Wishart) were applied in order to see possible differences among SAR observables in terms of information that they contain and their usefulness in classifying particular land cover type. Another important issue is labeling the clusters, and this work is carried out by decision tree classification that uses knowledge based approach. This classifier is implemented by scrupulous knowledge of data obtained by empirical evidence and their experimental validation. It has been demonstrated quantitatively that standard polarimetric parameters such as polarized backscatter coefficients (linear, circular

Received 14 January 2011, Accepted 13 April 2011, Scheduled 29 April 2011
Corresponding author: Dharmendra Singh (dharmfec@gmail.com).

and linear 45°), co and cross-pol ratios for both linear and circular polarizations can be used as information bearing features for making decision boundaries. This forms the basis of discrimination between various classes in sequential format. The classification approach has been evaluated for fully polarimetric ALOS PALSAR L-band level 1.1 data. The classifier uses these data to classify individual pixel into one of the five categories: water, tall vegetation, short vegetation, urban and bare soil surface. The quantitative results shown by this classifier give classification accuracy of about 86%, which is better than other classification techniques.

1. INTRODUCTION

Quantitative assessment of land cover is required for every country in order to make proper planning against earth surface alteration, since land cover change is related to global change due to its interaction with climate, ecosystem process, bio-geochemical cycles, biodiversity and human activity. Remote sensing plays an important role in land cover classification due to availability of various SAR images through ENVISAT, ALOS PALSAR, RADARSAT, TerraSAR-X, TanDEM-X and many others. Classifying remotely sensed data into a thematic map is a very challenging task because it depends upon many factors. The factors, such as complexity of the landscape in a study area, selected remotely sensed data, adopted image processing and classification approaches, may affect the success of a classification.

Since 1980's, radar polarimetry, i.e., the utilization of complete electromagnetic vector wave information, has been gaining more and more recognition from many researchers. Since then, radar polarimetry has been used in conjunction with remote sensing, and splendid results have been achieved. Cloud and Pottier [1, 2] made important contribution in the field of target decomposition by introducing the concept of Anisotropy, Entropy and alpha (α), and these parameters have become the standard tool for target characterization and have been used as the basis for the development of new classification methods introduced for the analysis of polarimetric data. Many approaches for target decomposition and land cover classification have been proposed notably by Pottier [3], Lee et al. [4, 7, 8], Krogager [5], Cameron and Leung [6], Ferro-Famil et al. [9], Ouarzeddine and Souissi [10], Fang et al. [11], Park et al. [12], Praks et al. [13]. There are mainly two groups of classification techniques — supervised and unsupervised. Although much research has been done in the field of SAR image classification, there are certain limitations in each classification technique due to their problem specific nature. For

example, besides being widely applicable the major disadvantage of supervised classification technique is that it is a single discriminative classifier which is applied to the individual pixel level or image objects (group of adjacent, similar pixels). If during training process any pixel is unidentified then supervised classifier cannot assign it to any class. Also supervised classifier finds it unable to recognize and represent unique categories not represented in training data. Unsupervised classification also suffers from certain limitations and disadvantages. One of the major disadvantages of unsupervised classification is that natural grouping obtained as a result of iterations in classifier does not necessarily correspond nicely to desired informational classes, and analyst has limited control over the classes chosen by the classification process [14, 15]. Thus in order to achieve more accurate results for land cover classification it is advantageous to opt for more advance classifier. In recent years the use of decision tree classifier for land cover classification of remotely sensed data has been increased considerably. Previous researches show that decision tree algorithms consistently outperform supervised classification techniques [15–17]. Decision tree classification is computationally efficient algorithm. The other advantages which decision tree classifier offers include its flexibility, simplicity, ability to handle noisy and missing data, lack of dependence on probability distribution function of data [16, 18]. Therefore in this paper the effort is mainly focused on classifying SAR image using knowledge based decision tree approach, in which the generic characteristics of various land covers are derived by backscattering coefficients for various polarizations. The decision rules for the classification were selected quantitatively on the basis of empirical evidence and experimental investigation. Another objective of this paper is to improve our understanding about supervised classification to see, how they interact with training data, and how they affect cluster labeling for land cover classification, if input parameters are SAR observables obtained by decomposition methods. In the present paper three supervised classification methods, namely minimum distance, maximum likelihood and parallelepiped, are used for classification based on Pauli decomposition and eigenvalue decomposition. The parameters obtained by eigenvalue decomposition, i.e., entropy, alpha and anisotropy are also taken as input parameters for Wishart classifier which is a state of art method used more often for classification.

A substantial amount of research has been carried out showing that fully polarimetric SAR systems are more capable in discriminating different land covers than single or dual polarimetric SAR data [19–21]. Thus in this paper ALOS-PALSAR level 1.1 fully polarimetric data has been used for study.

2. MATERIALS AND METHODS

2.1. Study Area

The study area has centre latitude 29.61380° and longitude 78.0086730°. It covers Roorkee, Laksar, Bijnor regions. Figure 1 shows study area on the map of India and on topographic map. Roorkee, a city in the state of Uttarakhand within India, is located at 29° 51′N, 77° 53′E on the south bank of Solani river. The upper Ganga canal, which runs from north to south, adds beauty to the city and divides the city in two distinct parts. Laksar is one of the three tehsils in Haridwar district in the state of Uttarakhand. Bijnor is a small town in the state of Uttar Pradesh. The elevation of study area ranges from 245.5 to 289.9 m. The study area has been chosen for its varied landscapes: water (source: Ganga canal, Solani river, a non perennial river, the most part of which remains dry in summers), urban (source: Roorkee, Laksar, Bijnor), tall vegetation (source: dense tree cover in city Roorkee), bare soil, short vegetation (source: crop land and grass land).

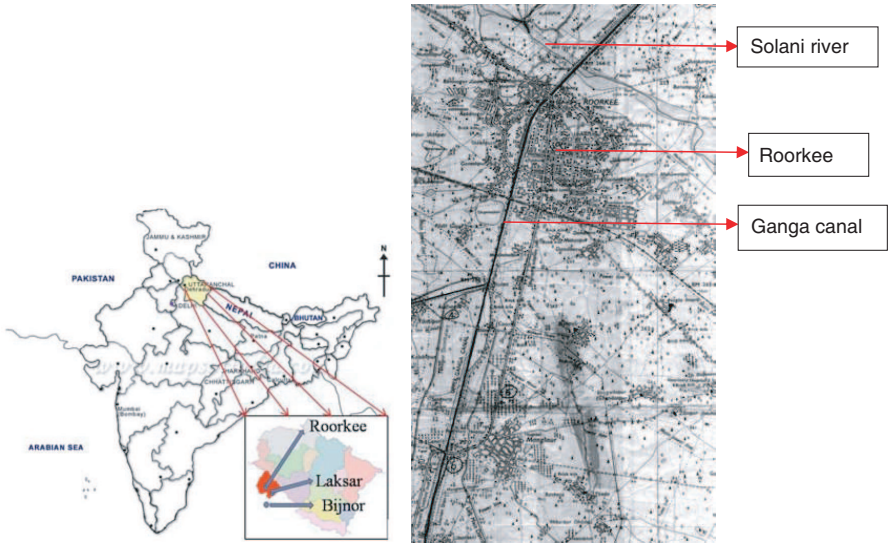


Figure 1. Location of study area in (a) map of India, (b) topographic map of city Roorkee.

2.2. Data Used

Two sets of ALOS PALSAR polarimetric data taken on 6 April, 2009 were used in the study. The first set of data (Data ID-PASL1100904061711181001150003, center latitude: 29.6138°, center longitude: 78.00867°) were used for development and testing of the proposed algorithm, and the second set of data (Data ID-PASL1109040617112609081100 63, center latitude: 30.109035°, center longitude: 77.892747°) were used for validation of algorithm. The data have four different modes: HH, HV, VH and VV polarizations. The ALOS PALSAR product is level 1.1 data in VEXCEL format, which are single look complex data on slant range. The product has single number of looks in range and azimuth. The default off nadir angle for polarimetric acquisition mode is 21.5°. Figure 2 shows colour coded processed image containing HH as red, HV as green and VV as blue.

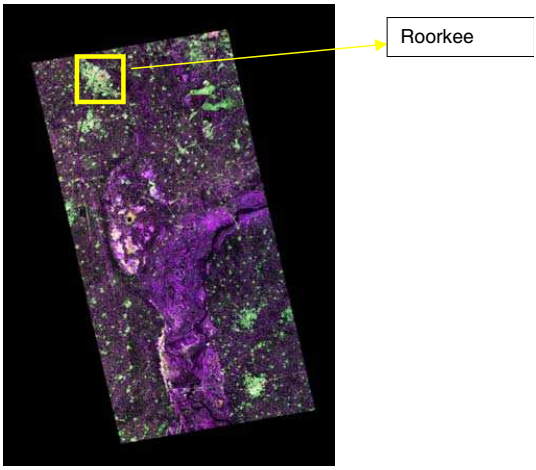


Figure 2. Colour composite image HH-HV-VV (HH — red, HV — green, VV — blue) to be used for classification for test data.

Table 1. Ground truth survey points.

Class	Training samples	Test samples
Water	64	302
Urban	67	265
Tall vegetation	19	70
Short vegetation	37	137
Bare soil	24	66

Extensive ground truth survey was performed over the whole region. Around 211 ground **control** points (GCP) were collected for training and 840 for testing the accuracy of classification map. Table 1 presents the training and control samples based on ground truth data. Based on ground truth information, five classes were identified: water (including wetland also), urban, short vegetation (cropland, grass land, shrubs etc.), tall vegetation and bare soil surface.

2.3. Pre-processing

The preprocessing steps are shown in Figure 3. The ALOS PALSAR L-1.1 data were first imported into SARSCAPE using ENVI 4.3 software. These SLC (single look complex) files were calibrated using defined polarimetric calibration matrices. Polarimetric calibration minimizes the impact of non ideal behavior of a full-polarimetric SAR acquisition system in order to obtain an estimate of the scattering matrix of the imaged objects as accurate as possible from their available measurement. The use of SAR data in classification requires that a radar adaptive filter is applied for speckle removal. The adaptive filter Wishart Gamma Map was selected due to its ability of preserving polarimetric information and the statistical properties between channels [22]. For eigenvalue decomposition boxcar filter was used. The size of processing window of filter is another important factor for classification of SAR images, since its size largely depends on the characteristics of target. Larger processing windows produce low resolution images due to amalgamation of various textural features, while using smaller windows the pixels in

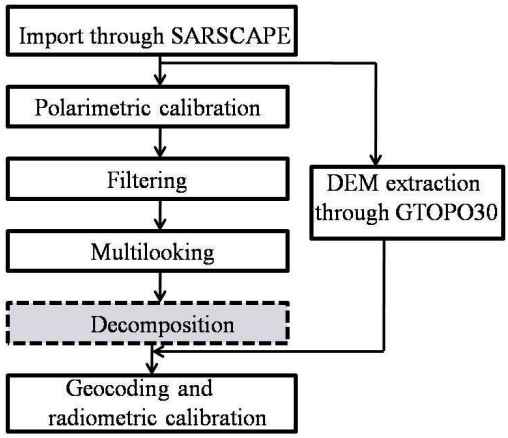


Figure 3. Flow chart for data pre-processing.

small proximity are considered, which identify the compact classes better. Also, for very small processing window second order statistics do not abide by averaged quantities of coherency matrix used in decomposition [23]. Since we are dealing with only backscattering coefficients of different polarizations (which are obtained by scattering matrix not by coherency matrix), we can select small window for accurate localization of boundaries. We have chosen 5×5 processing window for our site, which is neither very small nor very large and a good choice for classification. Due to different resolutions in range and azimuth direction multilooking was performed to improve the radiometric resolution. For ALOS PALSAR data multilooked factor was selected as 7, in order to avoid oversampling effect in geocoding. The digital elevation model was extracted using technique GTOPO 30 for terrain correction prior to geocoding. Then nearest neighbor resampling method was applied to data for radiometric calibration. In Figure 3 decomposition step is highlighted because this step is required only for H/A/alpha decomposition not for decision tree classification. It is also worth mentioning that in our whole study we have considered flat terrain with slope of around 4 degree for which any change in incidence angle across a distributed target is neglected because the incidence angle depends only on reference ellipsoid and varies from about 19.2° to 26.3° at far range.

2.4. Methodology

2.4.1. Target Decomposition Theorem

The decomposition in radar polarimetry provides a way for interpretation and optimum utilization of polarimetric scattering data by expressing the average mechanisms as the sum of independent elements. This leads to association of physical mechanism with each independent component having physical constraints such as the average target being invariant to changes in wave polarization basis [24]. Thus, decomposition technique manipulates the elements of scattering matrix or their second order statistics with the objective to provide more descriptive and discriminative target parameters, which have influential significance in various applications of radar polarimetry [1].

Target decomposition theorems were first formalized by Huynen but have their origin in the research work of Chandrasekhar on light scattering by small anisotropic particles [24]. At present, two theories of target decomposition can be distinguished: coherent target decomposition (CTD) and incoherent target decomposition (ICTD).

a. Coherent Target Decomposition (CTD): CTD deals with decomposition of scattering matrix, which characterizes the scattering

process from the target itself. This can happen only when incident and scattering waves are fully polarized. Consequently, CTD can only be employed to study coherent targets or point targets. The well-known Pauli decomposition is the basis of the coherency matrix formulation [4]. The Krogager decomposition decomposes a symmetric scattering matrix into three coherent components of sphere, diplane (dihedral), and helix targets [5]. Cameron classified a single target represented by a scattering matrix into many canonical scattering mechanisms that include trihedral, dihedral, dipole, 1/4 wave devices etc. [6]. In our study we have used only Pauli decomposition technique due to its simplicity.

a.1. Pauli Decomposition: The most common known and applied coherent decomposition technique is Pauli decomposition. The Pauli decomposition expresses the measured scattering matrix $[S]$ in the so-called Pauli basis [4]. The vectorization of $[S]$, carried out by using the Pauli matrices basis set, leads to the Pauli scattering vector or Pauli feature vector for the bi-static case.

b. Incoherent Target Decomposition (ICTD): The CTD approach shows inability in decomposing distributed targets. These scatterers can only be characterized statistically, due to the presence of speckle noise. To reduce speckle noise only second order polarimetric representations are required to analyze distributed scatterers. These second order descriptors are 3×3 Hermitian average covariance and the coherency matrices. ICTD deals with decomposition of these matrices. Consequently, ICTD also deals with partial polarized case. These types of decompositions are divided into two groups. First group is eigenvalue decomposition [1–3], where entropy, anisotropy and alpha parameters were introduced. Entropy is the measure of randomness of scattering, which can also be interpreted as degree of statistical disorder. Anisotropy can be defined as the normalized difference between the appearance probabilities of the second and third scattering components. From a practical point of view, the anisotropy can be employed as a source of discrimination only when entropy is greater than 0.7 because for lower entropies, the second and third eigenvalues are highly affected by noise. Consequently, the anisotropy is also very noisy [25]. The parameter α is an indicator of type of scattering and is called scattering mechanism. In general case it is often better to form a weighted average of the α parameters from the eigenvectors to obtain an average scattering mechanism. Such an average has been used for the interpretation of scattering by random particle volumes [26].

The second group is model based decomposition proposed by Freeman and Durden [27]. In this decomposition technique the covariance matrix was represented as the contribution of

three scattering mechanisms-single bounce scattering, double bounce scattering and volume scattering. By assuming reflection symmetry for the observation, this algorithm ignored the non-negligible power in off diagonal terms. This limitation was overcome by Yamaguchi et al. [28] who added forth term called helix scattering for representing off diagonal terms. Although these decomposition techniques are simple, they have certain limitations. One limitation of above two decomposition techniques is that both are based on thin cylinders, as primary scatterers for representation of vegetation canopy, which cannot accurately represent large leaves and thick branches [29]. Another limitation of these decomposition techniques is that volume scattering terms assume reflection symmetry condition for the observation [30]. Due to these limitations we did not use model based decomposition.

2.4.2. Knowledge Based Approach for Land Cover Classification (Decision Tree Approach)

Decision tree classifier, a machine learning algorithm, is a knowledge based data mining technique. It is an efficient tool for land cover classification. It is a hierarchal top-down approach, in which decision rules are defined by combination of several features, and a set of linear discriminate functions are applied at each test node, where a binary decision is made for splitting a complex decision into several simpler decisions so as to separate either one class or some of the classes from remaining classes. In this approach, the feature of data is a predictor variable (A variable analogous to the independent variable in linear regression and used for predicting the value of the target variable) whereas the class to be mapped is referred to as target variable. It performs binary recursive partitioning to allocate automatically maximum information carrying feature for the classification and discards remaining features at that transitional stage, thereby increases computational efficiency [31, 32].

3. CLASSIFICATION SCHEME

Classification is the task in which a set of given data elements (pixels) is assigned to some classes such that the cost of assigning data elements is minimum [37]. The intent of the classification process is to categorize all pixels in image into one of several land cover classes. These categorized data may then be used to produce thematic maps of the land covers present in an image. There are two main types of classification techniques: parametric and non-parametric. Parametric

classifiers are again of two types: supervised and unsupervised. Description of these techniques is beyond the scope of this paper, rather three supervised techniques (maximum likelihood, minimum distance and parallelepiped) are discussed in Appendix A. Decision trees, artificial neural networks, or support vector machines fall under the second category of classification, i.e., non parametric classification, which does not involve estimation of statistical parameters before classification [38–40].

3.1. Decision Tree Approach

Decision tree approach requires thorough knowledge of information bearing features and their physical understanding [33–35]. It has been already known that backscattering is a function of the electromagnetic wave parameters such as wave frequency, its polarization and its incidence angle, and it depends on the target characteristics such as surface geometry (size, shape, orientation distribution and spatial arrangement of objects), physical property (symmetry, non symmetry or irregularity of the target) and dielectric characteristics of the medium. The objective of the present work is to extract physical information from backscattering behavior of various objects. The analysis is based on polarized backscattering coefficients measured at HH-, VV-, HV-, RR- (circular copolar), RL- (circular cross polar), **45C- (45° co-polar**, σ°_{45XX} **and** σ°_{45YY}**), 45X- (45° cross polar** σ°_{45XY}**) and cross-pol** $\sigma^{\circ}_{HV}/\sigma^{\circ}_{VV}$ **and** $\sigma^{\circ}_{RR}/\sigma^{\circ}_{RL}$. These standard polarimetric features act as our information bearing features for predicting the nature of target.

Above mentioned all polarized backscatter coefficients were calculated for each class such as urban, water, short vegetation, tall vegetation and bare soil covered by GCPs. Since each class represents specific scattering property, decision boundaries are made based on knowledge acquired experimentally by the analysis of scattering behavior of each surface types. In order to make decision boundaries for separation of various classes, the concept of feature separation has been used [36]. The measure, called separability index for class pair separation, is defined as

$$S_{ij} = \frac{|\mu_i - \mu_j|}{\sigma_i + \sigma_j} \tag{1}$$

where μ and σ are mean values and standard deviations, respectively of classes i and j for a particular feature. Two classes will be well separated by a particular feature if the difference between the mean values of two classes is large as compared to the sum of standard deviations of those classes. In particular, S_{ij} lies in between 0.8 to

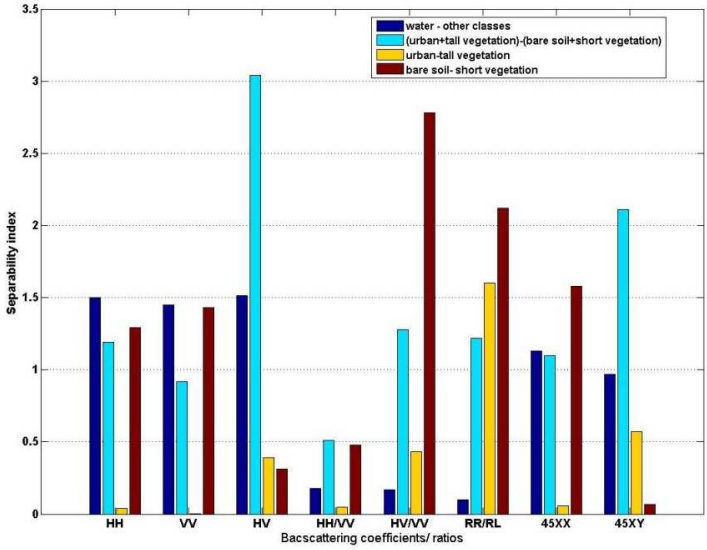


Figure 4. Separability index for class pair separation by various features (backscattering coefficients and backscattering ratios, HH refers to σ°_{HH}, HH/VV refers to $\sigma^{\circ}_{HH}/\sigma^{\circ}_{VV}$ and so on).

1.5, which indicates an authentic feature for separation of two classes i and j while S_{ij} greater than 2 represents the feature for nearly complete class pair separation. Figure 4 shows class separability index for various features namely, σ°_{HH}, σ°_{VV}, σ°_{HV}, $\sigma^{\circ}_{HH}/\sigma^{\circ}_{VV}$, $\sigma^{\circ}_{HV}/\sigma^{\circ}_{VV}$, $\sigma^{\circ}_{RR}/\sigma^{\circ}_{RL}$, σ°_{45XX} and σ°_{45XY} designated in Figure 4 as HH, VV, HV, HH/VV, HV/VV, RR/RL, $45XX$ and $45XY$ respectively. Figure 4 shows that $\sigma^{\circ}_{HH}/\sigma^{\circ}_{VV}$ is the worst feature for separation of all the defined class pairs.

For applying the concept of feature separation four different combinations of class pairs have been chosen. These combinations are i) water and others (urban, tall vegetation, bare soil and short vegetation), ii) (urban plus tall vegetation) and (bare soil plus short vegetation), iii) urban and tall vegetation, and iv) bare soil and short vegetation. The criterion for selecting the above defined class pairs is defined below-

i. Based on previous studies we know that water class shows very low reflection coefficient (backscattering coefficient) with respect to other classes [23]. So water was selected as one class while other classes, i.e., urban, tall vegetation, bare soil and short vegetation, were considered together as another single class. Now separability index was calculated for all the features in order to separate these

two class pairs. The useful features ($S_{ij} > 0.8$) for separating class water from other classes were σ°_{HH}, σ°_{VV}, σ°_{HV}, $\sigma^\circ_{HV}/\sigma^\circ_{VV}$, $\sigma^\circ_{RR}/\sigma^\circ_{RL}$ and σ°_{45XY}.

ii. Now four classes were left to deal with. It was observed that classes of tall vegetation and urban show nearly the same scattering characteristics for almost all the features due to similarity in their geometrical structure, so urban and tall vegetation were taken as one class, and bare soil and short vegetation as the other class. The separability index for separation of these two class pairs was greater than 0.8 for features σ°_{HH}, σ°_{VV}, σ°_{HV}, $\sigma^\circ_{HV}/\sigma^\circ_{VV}$, $\sigma^\circ_{RR}/\sigma^\circ_{RL}$, σ°_{45XX} and σ°_{45XY}. Out of these features perfect class pair separation ($S_{ij} > 2$) was shown by features σ°_{HV} and σ°_{45XY}.

iii. For the separation of classes urban and tall vegetation which were taken together as one class in step-ii, separability index was calculated which indicates that only one feature $\sigma^\circ_{RR}/\sigma^\circ_{RL}$ is able to segregate these two classes.

iv. In order to separate classes bare soil and short vegetation separability index was obtained, based on which reliable segregation was shown by σ°_{HH}, σ°_{VV}, $\sigma^\circ_{HV}/\sigma^\circ_{VV}$, $\sigma^\circ_{RR}/\sigma^\circ_{RL}$ and σ°_{45XX}. Out of these features perfect class pair separation ($S_{ij} > 2$) was shown by features $\sigma^\circ_{HV}/\sigma^\circ_{VV}$ and $\sigma^\circ_{RR}/\sigma^\circ_{RL}$.

The classification scheme for decision tree classification is shown in Figure 5. As discussed earlier we have used backscatter coefficients for various polarizations as our decision making criterion. It is known that SAR signals are sensitive to size, shape and orientation of targets. Therefore, we can predict the nature of landscapes by visualizing their scattering behavior, which can be better explained by backscattering coefficients of various targets at different polarizations (linear and circular). The main advantage of using polarized backscatter coefficient as information bearing feature is that it will reduce the need of 'a priori' knowledge about the test site.

The algorithm starts with discrimination between water and other classes. Based on empirical evidence and experimental validation, decision boundaries are created. It was observed by backscattering behavior that out of several features selected on the basis of separability index only σ°_{HV} and σ°_{45XY} can uniquely separate water bodies from other classes. For our site σ°_{HV} is less than $-30\,\mathrm{dB}$, and σ°_{45XY} is less than $-25\,\mathrm{dB}$. All the areas with σ°_{HV} greater than $-18\,\mathrm{dB}$ [41], σ°_{45XY} greater than $-20\,\mathrm{dB}$ and $\sigma^\circ_{HV}/\sigma^\circ_{VV}$ greater than $-11\,\mathrm{dB}$ [42] are classified as tall vegetation as well as urban. These two classes are separated on the basis of $\sigma^\circ_{RR}/\sigma^\circ_{RL}$ which is negative for urban and positive for tall vegetation. It has been verified that for surface scattering, e.g., for bare soil, σ°_{RR} is appreciably less than σ°_{RL} [43], and

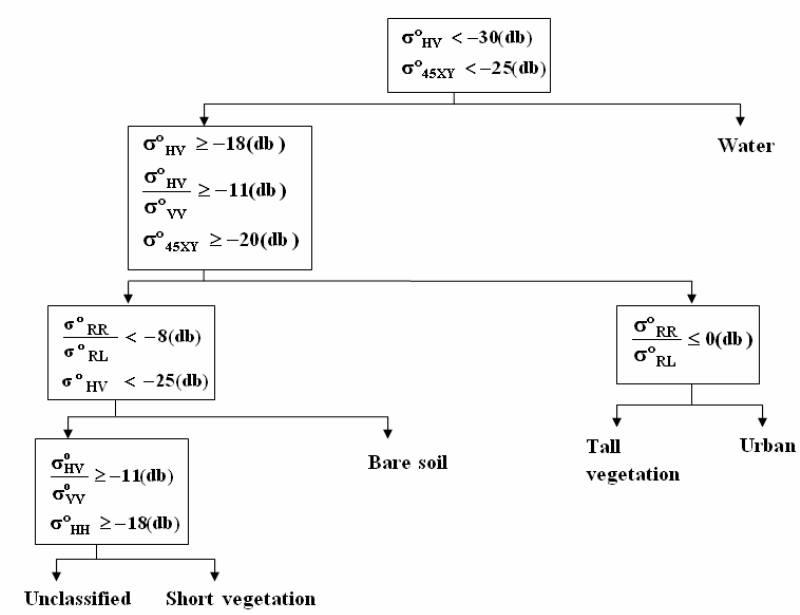


Figure 5. Algorithm for decision tree classification.

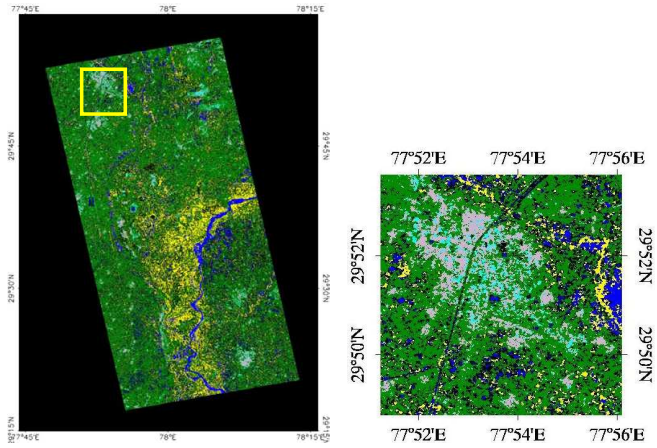
σ°_{HH} is less than σ°_{VV} [39]. The factor $\sigma^{\circ}_{HH}/\sigma^{\circ}_{VV}$ is not included due to very low separability index. In our site, $\sigma^{\circ}_{RR}/\sigma^{\circ}_{RL}$ is less than -8 dB, and σ°_{HV} is less than -25 dB for bare soil surface. This allows the segregation of bare soil surface from other classes. As we have already known, for vegetation $\sigma^{\circ}_{HV}/\sigma^{\circ}_{VV}$ is greater than -11 dB. Also we have found from backscattering behavior that σ°_{HH} is greater than -18 dB for short vegetation. The pixels that do not satisfy above criteria are unclassified.

4. RESULTS AND DISCUSSION

This section presents the results for all the classifications, i.e., decision tree classification, supervised classifications (minimum distance, maximum likelihood, parallelepiped) based on Pauli decomposition and eigenvalue decomposition and Wishart classification based on Eigenvalue decomposition. The result of classification algorithm was calculated using confusion matrix (or error matrix), which compares the classification result with ground truth information and reports overall accuracy, kappa coefficient, producer accuracy and user accuracy. All the results were obtained by SARSCAPE processing tools.

4.1. Decision Tree Based Classification

According to the classification scheme shown in Figure 5, the algorithm developed by ENVI 4.3 was run on pixel by pixel basis. The result of this classification scheme is shown in Figures 6(a) and 6(b) in which classification result can clearly be visualized. These figures show that most of the pixels belonging to specific field are classified as the same category. We observe very coherent result for each of the individual fields.



(a) (b)

Figure 6. (a) Result of decision tree classifier of test site. (b) Highlighted part of image (a) showing region Roorkee: (Water-blue, urban-pink, bare soil-yellow, short vegetation-green, tall vegetation-cyan).

The confusion matrix for accuracy assessment is shown in Table 2. The overall classification accuracy is estimated as 86% and kappa coefficient as 0.82. Producer accuracy and user accuracy estimate relative to decision tree classification are shown in Table 3.

Table 2. Pixel assignment of various classes shown by confusion matrix relative to decision tree classification.

Class	Water	Urban	Tall veg.	Short veg.	Bare soil	Total
Water	271	0	0	14	20	305
Urban	0	203	2	0	0	205
Tall veg.	0	2	67	0	0	69
Short veg.	1	54	1	57	0	113
Bare soil	3	0	0	10	46	59
Total	275	259	70	81	66	

Table 3. User accuracy and producer accuracy (in percent) estimates relative to decision tree classification.

Class	User accuracy (%)	Producer accuracy (%)
Water	88.85	98.55
Urban	99.02	78.38
Tall vegetation	97.10	95.71
Short vegetation	50.04	70.37
Bare soil	77.97	69.70

4.2. Classification Based on Pauli Decomposition

The classification has been carried out based on Pauli decomposition (Section 2.4.1) and producer accuracy estimates for classification tests based on Pauli decomposition are shown in Table 4.

The classification accuracy is maximum for maximum likelihood classifier with overall accuracy of 71.54% and kappa coefficient of 0.6203, while minimum classification accuracy is shown by parallelepiped classifier with overall accuracy of 30.5952% and kappa coefficient of 0.2075. The overall classification accuracy for minimum distance classifier lies in between that of maximum likelihood and parallelepiped classifier which is 66.4286% with kappa coefficient of 0.5469.

Some limits are shown by parallelepiped classifier in classifying classes "water" and "urban". This classification technique completely fails in recognizing training pixels related to class "water" and classifies class "urban" with producer accuracy of only 3.77% because the class "tall vegetation" is misclassified.

Maximum likelihood classifier identifies all the training pixels more accurately than others and classifies all land cover types with satisfactory performance indices, since each class has quite good producer accuracy. Minimum distance classifier also shows almost the same results. It classifies class "water" more accurately than maximum likelihood classifier (see Table 4).

Table 4. Producer accuracy estimate (in percent) for classification test based on Pauli decomposition.

Classification	Water	Urban	Tall veg.	Short veg.	Baresoil
Parallelepiped	2.5	3.77	98.57	93.43	75.76
Minimum distance	78.81	68.30	48.57	54.01	46.97
Maximum likelihood	59.60	91.32	34.29	73.72	81.82

4.3. Classification Based on Eigenvalue Decomposition

The classification tests based on eigenvalue decomposition were performed by two methods: supervised (Minimum distance, Maximum likelihood, Parallelepiped) and unsupervised (Wishart). The Wishart classification was performed by PolSARpro software freely available by ESA (European Space Agency) website. The producer and user accuracy estimates relative to Wishart classifier for H/A/α are shown in Table 5. The overall classification accuracy was obtained as 43.35%, and kappa coefficient was 0.2731. The supervised classification was performed by performing two series of experiments. First series of experiment was performed using entropy (H) and alpha (α) as an input to classifier. The producer accuracy estimates for classification tests based on H/α are shown in Table 6. The second series of experiment was performed using all three eigenvalue parameters, i.e., entropy (H), alpha (α) and anisotropy (A).

Table 5. User accuracy and producer accuracy (in percent) estimates relative to H/A/alpha based on Wishart classification.

Class	User accuracy (%)	Producer accuracy (%)
Water	38.57	21.77
Urban	86.02	52.29
Tall vegetation	8.70	10.53
Short vegetation	41.75	76.79
Bare soil	12.64	45.83

Table 6. Producer accuracy estimate (in percent) for classification test based on H/α.

Classification	Water	Urban	Tall veg.	Short veg.	Baresoil
Parallelepiped	0.23	53.96	30.00	78.10	86.36
Minimum dis.	17.88	43.02	80.00	37.96	93.94
Maximum like.	11.26	56.98	55.71	52.55	93.94

In both series of tests, overall accuracy remains low (around 40%). The kappa coefficient, which is an indicator of performance of classification lies below 0.4 in all the cases. The classification tests are not sensitive to additional anisotropy information. Rather, after including anisotropy as an input parameter to classifiers, overall accuracy actually decreases.

The parallelepiped classification based on parameters H and α gives poor results in recognizing class "water" from training pixels. Class "water" is also poorly classified by minimum distance and maximum likelihood classifier with producer's accuracy of 17.88% and 11.26%, respectively. "Bare soil" is classified perfectly by all the classifiers (Table 6). Class "urban" is well defined in parallelepiped classification for both series of tests. In minimum distance classification "urban" and "tall vegetation" classes are mingled with each other.

For the second series of test based on parameters H, A, and α, parallelepiped classifier again performs poorly in classifying class "water". The producer's accuracy for class water by parallelepiped classification is only 0.33%. Class "bare soil" again is classified correctly by all three classifiers. The confusion matrix for the second series of tests is shown in Table 7.

Table 7. Producer accuracy estimate (in percent) for classification test based on H/A/α.

Classification	Water	Urban	Tall veg.	Short veg.	Baresoil
Parallelepiped	0.33	56.98	20.00	69.34	90.91
Minimum distance	17.88	43.02	80.00	37.96	92.42
Maximum likelihood	11.92	55.09	45.71	46.72	87.08

Table 8. Overall accuracy (O.A.) and kappa coefficient estimates for all the features obtained by decomposition techniques.

Features	Maximum likelihood		Minimum distance		Parallelepiped	
	O.A.	Kappa	O.A.	Kappa	O.A.	Kappa
H/Alpha	42.61	0.2643	40.23	0.2444	39.04	0.2325
H/A/Alpha	40.00	0.2771	40.12	0.2431	38.21	0.2186
Pauli decomposition	71.54	0.6203	66.42	0.5469	30.59	0.2075

The classification results showing overall classification accuracy and kappa coefficient for different methods are summarized in Table 8. The results show that decision tree classifier performs better than all three parametric supervised classification techniques (maximum likelihood, minimum distance and parallelepiped) and Wishart classification, because this classifier is implemented with thorough knowledge of data obtained by empirical evidence and experimental validation which does not require to make any prior

assumption about the frequency distribution or other statistical properties of data.

In non-parametric case, maximum likelihood classifier performs better than minimum distance and parallelepiped classifiers. The possible reason is that the maximum likelihood classifier is based on Bayesian probability theory which uses second order statistics unlike minimum distance and parallelepiped classifiers which adopt first order statistics. This maximum likelihood classifier assumes that input data (training data) is normally distributed and independent.

The classification tests based on H/A/alpha and H/alpha give not so good results for all three classification algorithms. The possible reason of the poor result is the use of boxcar filter for speckle removal. The boxcar filter causes sharp edges to be blurred, and transforms point scatterer to spread target due to over-filtering. Wishart classification also does not give good results.

4.4. Algorithm Validation

The proposed algorithm was validated on another PALSAR data (data ID-PASL1109040617112609081100 63) of date 6th April, 2009 for which overall classification accuracy was obtained as 77.69% with kappa coefficient of 0.7156. It was found that producer accuracy for water was 92.75%, tall vegetation - 52.44%, short vegetation - 77.53%, urban - 84.80% and bare soil - 80.33%. The results are shown in Figure 7.

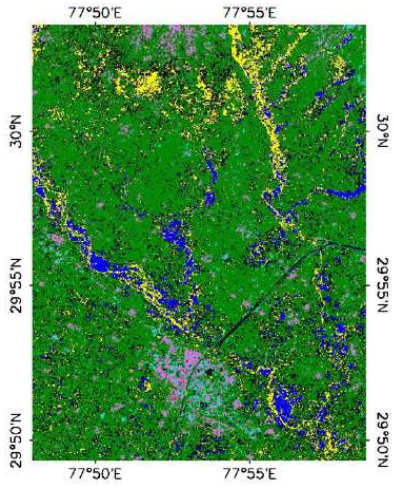


Figure 7. Result of decision tree classifier validated on second data. (Water-blue, urban-pink, bare soil-yellow, short vegetation-green, tall vegetation-cyan).

5. CONCLUSION

A machine learning decision tree classifier is presented in this paper for land cover classification. This classification scheme is applied to fully polarimetric ALOS PALSAR L-1.1 data. Five classes were identified: water, urban, tall vegetation, short vegetation, and bare soil. Polarized backscatter coefficients, co- and cross-pol ratios were analyzed, which confirm the work of previous research. The classifier was developed with expert knowledge based on empirical evidence and experimental validation, with each class having its own classification rule due to its specific scattering behaviour. The result shows fairly good classification accuracy for each class. Dependence on physical principles, which does not require any prior assumption, makes this method simple, robust and widely applicable. We have also performed land cover classification by three supervised classification techniques, namely, maximum likelihood, minimum distance and parallelepiped, based on SAR observables obtained by Pauli decomposition and Eigenvalue decomposition techniques and Wishart classification based on SAR observables obtained by eigenvalue decomposition. The decision tree classifier recognizes all land cover types more accurately from training pixel than parallelepiped, minimum distance, maximum likelihood and Wishart classifiers. In this paper, we have tested the proposed algorithm on L-band data, and it works well for this band. In order to apply to different band data we have to check parameter values again for training decision tree because it is well-known that radar response to various land covers is polarization and frequency dependent.

ACKNOWLEDGMENT

The authors are thankful to Department of Science and Technology, India and Japan Science and Technology, Japan for providing funds to carry out this work. The authors are also thankful to JAXA, Japan for providing ALOS-PALSAR data.

APPENDIX A. SUPERVISED CLASSIFICATION SCHEMES

Supervised classification involves using a priori knowledge of data to "train" computer software to identify categories in an image [39]. The supervised approach to pixel labeling requires the user to select representative training data for each predefined number of classes. The main supervised classification techniques are defined below.

A.1. Parallelepiped Classification

The parallelepiped classifier characterizes each class by range of expected values on each band. The range is defined by maximum and minimum pixel values in a given class or alternatively by a certain number of standard deviations on either side of mean of training data for a given class. These decision boundaries form n-dimensional parallelepiped. If a pixel value lies above the low threshold and below the high threshold, it is assigned to that class. If the pixel value falls in multiple classes, the pixel is assigned to the last class matched or to overlap class. If the pixel does not fall within any of the parallelepiped classes it is designated as unclassified or null class. The advantage of this technique is that it is very simple and easy to implement whereas performance wise it is poor, and for correlated data there can be overlap of the parallelepipeds since their sides are parallel to the spectral axes. Consequently, there are some data that cannot be separated.

A.2. Minimum Distance Classification

The decision rule adopted by the minimum distance classifier to determine a pixel's label is the minimum distance between the pixel and the class centers (mean), measured either by the Euclidean distance or the Mahalanobis generalized distance. Classification is then performed by placing a pixel in the class of the nearest mean [38]. The advantage of this classifier is that it not only is a mathematically simple and computationally efficient technique, but also provides better accuracy than maximum likelihood procedure, in the case when the number of training samples is limited. But the shortcoming is that by characterizing each class by its mean reflectance only, minimum distance classifier has no knowledge of the fact that some classes are naturally more variable than others, which consecutively can lead to misclassification.

A.3. Maximum Likelihood Classification

The MLC procedure is based on Bayesian probability theory. Decision rule is decided by calculating mean and standard deviation of each training set and deriving probability density function from mean and standard deviation for computing probability of each pixel belonging to each class. The classifier then assigns pixel to the class for which the probability is the highest. It yields higher accuracy than other classifiers. It has some demerits like: (i) It is computationally intensive and time consuming technique; (ii) Each data sample has to be tested

against all classes in a classification, which leads to relative degree of inefficiency; (iii) With a fixed relatively small size training set the classification accuracy may actually decrease when the number of features are increased.

REFERENCES

1. Cloude, S. R. and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 34, No. 2, 498–518, 1996.
2. Cloude, S. R. and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 3, No. 1, 68–78, 1997.
3. Pottier, E., "Radar target decomposition theorems and unsupervised classification of full polarimetric SAR data," *Proceedings of IEEE Geoscience and Remote Sensing Symposium*, 1139–1141, 1994.
4. Lee, J. S., M. R. Grunes, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on the complex Wishart distribution," *International Journal of Remote Sensing*, Vol. 15, No. 11, 2299–2311, 1994.
5. Krogager, E., "A new decomposition of the radar target scattering matrix," *Electronics Letter*, Vol. 26, No. 18, 1525–1526, 1990.
6. Cameron, W. L. and L. K. Leung, "Feature motivated polarization scattering matrix decomposition," *Proceedings of IEEE International Radar Conference*, Arlington, VA, May 7–10, 1990.
7. Lee, J. S., M. R. Grunes, T. L. Ainsworth, et al., "Unsupervised classification of polarimetric SAR images by applying target decomposition and complex Wishart distribution," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, No. 5, 2249–2258, 1999.
8. Lee, J. S., M. R. Grunes, E. Pottier, et al., "Unsupervised terrain classification preserving polarimetric scattering characteristics," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 42, No. 4, 722–731, 2004.
9. Ferro-Famil, L., E. Pottier, and J. S. Lee, "Unsupervised classification of multi-frequency and fully polarimetric SAR images based on the H/A/Alpha-Wishart classifier," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 11, 2332–2342, 2001.

10. Ouarzeddine, M. and B. Souissi, "Unsupervised classification using wishart classifier," *USTHB, F.E.I.*, BP N° 32, EI Alia Bab Ezzouar, Alger, 2007.
11. Fang, C., H. Wen, and W. Yirong, "An improved Cloude-Pottier decomposition using H/α/span and complex Wishart classifier for polarimetric SAR classification," *International Conference on Radar*, 2006.
12. Park, S. E. and W. M. Moon, "Unsupervised classification of scattering mechanisms in polarimetric SAR data using fuzzy logic in entropy and alpha plane," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 45, No. 8, 2652–2664, 2007.
13. Praks, J., E. C. Koeniguer, et al., "Alternatives to target entropy and alpha angle in SAR polarimetry," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 47, No. 7, 2262–2274, 2009.
14. Enderle, D. I. M. and R. C. Weih Jr., "Integrating supervised and unsupervised classification methods to develop a more accurate land cover classification," *Journal of the Arkansas Academy of Science*, Vol. 59, 65–73, 2005.
15. Hui, Y., et al., "Extracting wetland using decision tree classification," *Proceedings of the 8th WSEAS International Conference on Applied Computer and Applied Computational Science*, 240–245, 2004.
16. Friedl, M. A. and C. E. Brodley, "Decision tree classification of land cover from remotely sensed data," *Remote Sensing of Environment*, Vol. 61, 399–409, 1997.
17. Pal, M. and P. M. Mather, "An assessment of the effectiveness of decision tree methods for land cover classification," *Remote Sensing of Environment*, Vol. 86, 554–565, 2003.
18. De'ath, G. and K. E. Fabricius, "Classification and regression trees: A powerful yet simple technique for ecological data analysis," *Ecology*, Vol. 81, No. 11, 3178–3192, 2000.
19. Ferro-Famil, L. and E. Pottier, "Dual frequency polarimetric SAR data classification and analysis," *Progress In Electromagnetics Research*, Vol. 31, 247–272, 2001.
20. Lee, J. S., M. R. Grunes, and E. Pottier, "Quantitative comparison of classification capability: Fully polarimetric versus dual and single-polarization SAR," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 11, 2001.
21. McNaim, H., J. Shang, X. Jiao, and C. Champagne, "The contribution of ALOS PALSAR multipolarization and polarimetric data to crop classification," *IEEE Transactions on Geoscience and Re-*

mote Sensing, Vol. 47, No. 12, 3981–3992, 2009.

22. Nezry, E. and F. Y. Simen, "On the preservation of polarimetric signatures and polarimetric texture signatures by fully polarimetric MAP filters," *Proceedings of IEEE Geoscience and Remote Sensing Symposium*, 1999.
23. Chamundeeswari, V. V., D. Singh, and K. Singh, "An adaptive method with integration of multi-wavelet based features for unsupervised classification of SAR images," *Journal of Geophysics Eng.*, Vol. 4, No. 4, 384–393, 2007.
24. Agrawal, A. P. and W. M. Boerner, "Redevelopment of Kennaugh's target characteristic polarization state theory using the polarization transformation ratio for the coherent case," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 27, 2–14, 1988.
25. Cloude, S. R., "Uniqueness of target decomposition theorems in radar polarimetry," *Direct and Inverse Methods in Radar Polarimetry*, W. M. Boerner et al. (eds.), Part 1, 267–296, Kluwer Academic Publishers, ISBN 0-7923-1498-0, NATO-ARW, 1992.
26. Cloude, S. R., E. Pottier, and W. M. Boerner, "Unsupervised image classification using the Entropy/Alpha/Anisotropy Method in radar polarimetry," *Proceedings JPL AIRSAR Symposium*, Pasadena, 2002.
27. Freeman, A. and S. L. Durden, "A three-component scattering model for polarimetric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 36, No. 3, 963–973, 1998.
28. Yamaguchi, Y., T. Moriyama, M. Ishido, and H. Yamada, "Four-component scattering model for polarimetric SAR image decomposition," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 43, No. 8, 1699–1706, 2005.
29. Arii, M., J. J. van Zyl, and Y. Kim, "A general characterization of polarimetric scattering from vegetation canopy," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, No. 9, 3349–3357, 2010.
30. Arii, M., J. J. van Zyl, and Y. Kim, "Adaptive model-based decomposition of polarimetric SAR covariance matrices," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 49, No. 3, 1104–1113, 2011.
31. Swain, P. H. and H. Hauska, "The decision tree classifier: Design and potential," *IEEE Transactions on Geoscience Electronics*, Vol. 15, 142–147, 1977.
32. Li, S., B. Zhang, L. Gao, and L. Zhang, "Classification of coastal

zone based on decision tree and PPI," *Proceedings of IEEE Geoscience and Remote Sensing Symposium*, 188–191, 2009.

33. Ferrazzoli, P., et al., "The potential of multi-frequency polarimetric SAR in assessing agricultural and arboreous biomass," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 35, 5–17, 1997.
34. Skriver, H., J. Dall, et al., "Agriculture classification using POL-SAR data," *Proceedings of the 2nd International Workshop on Applications of SAR Polarimetry and Polarimetric Interferometry*, 2005.
35. Pierce, L. E., F. T. Ulaby, K. Sarabandi, and M. C. Dobson, "Knowledge-based classification of polarimetric SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, 1081–1086, 1994.
36. Wu, F., C. Wang, H. Zhang, B. Zhang, and Y. Tang, "Rice crop monitoring in south China with RADARSAT-2 quad-polarization SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 8, 196–200, 2011.
37. Turkar, V. and Y. S. Rao, "Classification of polarimetric synthetic aperture radar images from SIR-C and ALOS PALSAR," *Proceedings of International Conference on Microwave 08*, 2008.
38. Richards, J. A., *Remote Sensing Digital Image Analysis*, Springer-Verlag, Berlin, 1999.
39. Lillesand, M. T. and R. W. Kiefer, *Remote Sensing and Image Interpretation*, 4th Edition, John Wiley & Sons, New York, 2000.
40. Tso, B. and P. Mather, *Classification Methods for Remotely Sensed Data*, 2nd edition, CRC Press, Taylor & Francis Group, LLC, 2009.
41. Ferrazzoli, P., L. Guerriero, and G. Schiavon, "Experimental and model investigation on radar classification capability," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, 960–968, 1999.
42. Dubois, P. C., J. Van Zyl, and T. Engman, "Measuring soil moisture with imaging radars," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 33, 915–926, 1995.
43. Baronti, S., F. Del Frate, P. Ferrazzoli, S. Paloscia, P. Pampaloni, and G. Schiavon, "SAR polarimetric features of agricultural areas," *International Journal of Remote Sensing*, Vol. 16, No. 14, 2639–2656, 1995.