POLARIMETRIC-SAR CLASSIFICATION USING FUZZY MAXIMUM LIKELIHOOD ESTIMATION CLUSTERING WITH CONSIDERATION OF COMPLEMENTARY INFORMATION BASED ON PHYSICAL POLARIMETRIC PARAMETERS, TARGET SCATTERING CHARACTERISTICS, AND SPATIAL CONTEXT

KATMOKO ARI SAMBODO'. ANIAT1 MURNI', RAT1II DEWANTI', AND MAHDI KARTASASMITA"

Abstract. This paper shows a study on an alternative method for unsupervised classi Ileal ion of polarimetric-Synthetic Aperture Radar (SAR) data. The first step was lo extract several main physical polarimetric parameters (polarization power, coherence, and phase difference) from polarimetric covariance matrix (or coherency matrix) and physical scattering characteristics of land use/cover based on polarimetric decomposition (Cloude decomposition model). In this paper, we found that these features have complementary information which can be integrated in order to improve the discrimination of different land use/cover types. Classification stage was performed using Fuzzy Maximum Likelihood Estimation (FMLE) clustering algorithm. FMLE algorithm allows for ellipsoidal clusters of arbitrary extent and is consequently more flexible than standard Fuzzy K-Means clustering algorithm. However, basic FMLE algorithm makes use exclusively the spectral (or intensity) properties of the individual pixel vectors and spatial-contextual information of the image was not taken into account. Hence, poor (noisy) classification result is usually obtained from SAR data due lo speckle noise. In this paper, we propose a modified FMLE which integrate basic FMLE clustering with spatial-contextual information by statistical analysis of local neighbourhoods. The effectiveness of the proposed method was demonstrated using E-SAR polarimetric data acquired on the area of Penajam, East Kalimantan, Indonesia. Results showed classified images improving land-cover discrimination performance, exhibiting homogeneous region, and preserving edge and other fine structures.

Keywords: Ctoude's polarimetric decomposition. FMLE clustering, polarimetric coherence, Polarimetric-SAR, unsupervised classification.

1. Introduction

Fully Polarimetric-SAR sensors are becoming more and more important in remote sensing applications due to: I) its all-weather, day and night operational capability; 2) its sensitivity of the polarization state of the backscattered wave to physical characteristics of the ground target (e.g. shape, size, orientation, surface roughness, moisture content, dielectric properties of the target) (Woodhouse, 2006; Tso and Mather, 2001; Bruzzone *el* a/., 2004). The utilization of multi-polarized wave in polarimetric-SAR system allows us to extract additional information which can be employed as classification features, thus giving better land use/cover classification results than single-channel single-polarization SAR data (Karathanassi and Dabboor, 2004; Woodhouse, 2006). For this reasons, in recent years, the remote sensing community has become increasingly interested in the use of polarimetric-SAR data for the production of high accuracy land-cover maps.

' Indonesian National Institute of Aeronautics and Space (LAPAN)
* University of Indonesia

Many algorithms have been proposed for supervised and unsupervised classification of polarimetric-SAR dala. In supervised classification approach, the choice of training areas which adequately represent the spectral characteristics of each class is important as Ihe quality of the training set has a profound effect on the validity of the result. This manual technique of finding and verifying training areas can be laborious, and particularly more complicated when ground truth (which often both cost and time consumption) or other priori information about the data is not available. On the other hand, unsupervised classification compensate for these deficiencies by finding an Underlying class structure automatically and then organizing the data into groups sharing similar characteristics (Canty, 2006). Unsupervised classification for polarimetric-SAR follows two major approaches. One is based on purely on statistical clustering of polarimetric-SAR data, and the other is based on the analysis of physical scattering properties. In former approach, the 3x3 complex covariancc matrix (or coherency matrix) formed from SAR polarimetry system measurements was assumed to have a multivariate complex Wishart distribution. Researchers use this distribution to derive distance measures for various clustering algorithm(Kersten *et al.,* 2005; Davidson *elai,* 2002).

The later approach used the inherent characteristics of polarimetric-SAR data and classified based on scattering mechanisms of the target scene. Fully polarimetric data provides unique possibility to separate scattering contributions of different nature, which can be associated to certain elementary scattering mechanisms (e.g. surface or single-bounce, double-bounce, and volume scattering). Several decomposition techniques have been proposed for extracting and identifying these valuable information. One method is based on polarimetric target decomposition theory proposed by Cloude and Pottier (1997), which is capable of covering whole range of scattering mechanisms and yields an' unsupervised classification scheme. The target's scattering mechanism can be parameterized by entropy //and alpha angle a which derived from the eigenvalue decomposition of the coherency matrix. The entropy *H* is a measure of randomness of scattering mechanisms, and the alpha angle a characterizes the scattering mechanisms. The H-a plane was divided into eight zones. The physical scattering characteristic associated with each zone provides information for terrain type assignment.

Additionally, several interesting combinations of these types of classification approaches have been found (Lee *el al.,* 1999-a: Kersten *et al..* 2005). Lee *et at.* (1999-a) developed some classification methods combining both the Cloude's polarimetric decomposition and clustering algorithm based on the complex Wishart distribution (often referred as complex Wishart classifier). These methods achieve preliminary classification sets using the H-a plane (8 zone), and thereafter iteratively classify these preliminary sets using Wishart classifier to make final classifications. Unsupervised Wishart H-a classification has been found to be applicable to land cover classification (Lumsdon, 2003), sea ice classification (Scheuchl, 2001), and forest classification (Ferro-Famile/o/.. 2005).

In general, acceptable classification results were obtained, w hereas in some cases, they also reported some limitations of these methods for further possibility to discriminate and classify into different object/ land cover types especially with similar scattering mechanism and often yield clusters (classes) whose physical meaning is uncertain. To overcome these problems, it is advisable to use the additional information which can be included as extension input features thus reduce inter-class ambiguity

and improve the classification performance. Although such additional information can be obtained from other data sources (such optical data, muIti-frequency radar data, geological maps, etc), the consideration of additional information which can be extracted directly from same polarimetric-SAR data but using different aspect would be meaningful (such image texture, context, structural relationships, etc). For example, in our previous publication (Sambodo *el al.*, 2007), we show that integration the combined features extracted from polarimetric decomposition and textural analysis with supervised neural network classifier successfully improve the classification results in a significant way. However, the aforementioned Wishart classifier use a 3x3 complex covariance matrix (or coherency matrix) form as an input feature, thus other additional features (which usually represent as various data types) can not be added into this input form.

Another limitation of these algorithms is that they performed on a pixel-by-pixel basis, i.e., each pixel is treated independently of its neighbours; spatial context is only indirectly considered during speckle filtering. The local neighbourhood does indeed have a significant influence on a pixel's class membership. When a certain region already has already been classified, with high confidence, as belonging to a single class, it becomes comparatively unlikely that a pixel in this region belongs to another class. The much more likely scenario is misestimation of its covariance matrix due to speckle noise, which usually produced very noisy classification results (often appear as "salt-and-pepper" effect even in homogeneous areas). Due to inherently high noise level of SAR data, the inclusion of local neighbourhoods in statistical decision about class membership is helpful to support homogeneous classification results (Canty, 2006).

In this paper, we propose an unsupervised classification method based on FMLE clustering algorithm that integrates complementary information of several polarimetric parameters and target scattering characteristic features, and spatial contextual information. Fuzzy classification techniques allow each pixel in the image to belong to more than one cluster according to its degree of membership in each clusters (Canty, 2006: Tso and Mather, 2001). Therefore, it is suitable tor classification of SAR data as the presence of speckle noise often causes many pixels in the data arc really ambiguous (i.e., imprecise, incomplete, and not totally reliable). A FMLE clustering has been chosen which it allows for hyperellipsoidal forms of the clusters and is consequently more flexible than standard fuzzy k-means (FKM) clustering (with the use of Euclidean distance, thus giving circular clusters) (Canty, 2006; Canty and Nielsen, 2004).

Further advantage is that other features can be easily added into FMLE clustering process by extending the dimension of the input data vectors. These properties enable us to combine the wide range of information (features) which can be derived from polarimetric-SAR data using different feature extraction methods. In our case, motivate by our previous publication (Sambodo *el a/.*, 2007), we will combine several main polarimetric parameters (polarization power, coherence, and phase difference) extracted from polarimetric covariance matrix and physical scattering characteristics of land use/cover based on Cloude's polarimetric decomposition. These features have complementary information which can be integrated in order to improve the discrim ination of different land use/cover types.

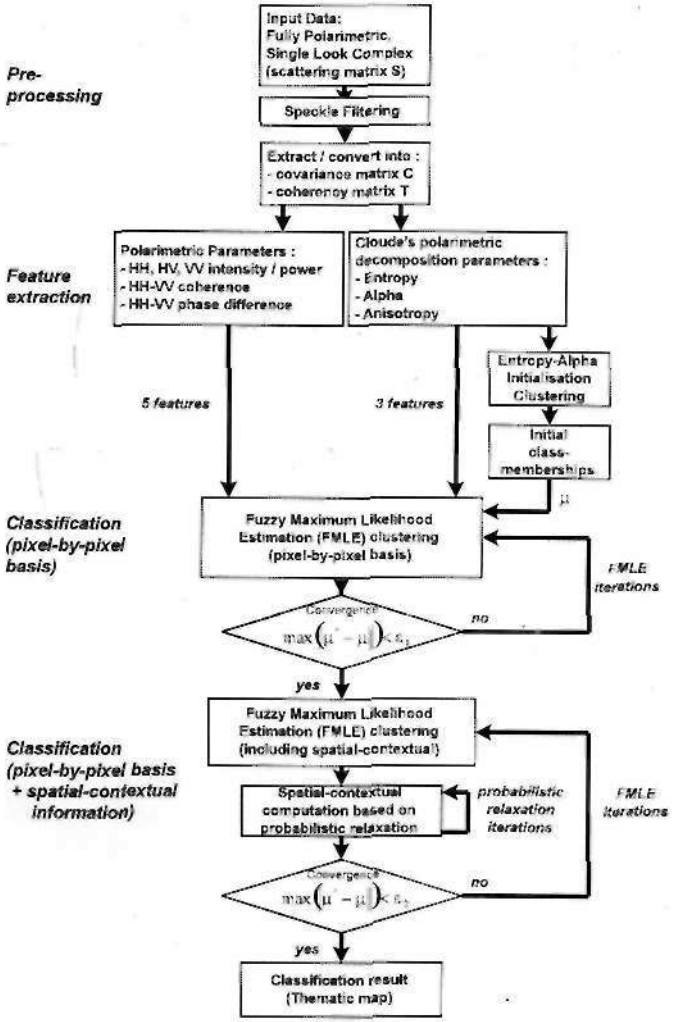


Figure 1. Block scheme of the proposed method.

However, the basic FMLE algorihtm is a pixel-by-pixel basis classifier. Thus, in order to exploit the spatial-contextual information, we investigate the possibility of using probabilistic relaxation schemes. It iteratively adjust some initial estimates of the class-membership probabilities by reference to the class-membership probabilities of pixels in its neighborhood. In this paper, we propose a contextual FMLE classification which integrate probabilistic relaxation scheme into FMLE clustering iterations in order to improve the estimation of clustering parameters themselves thus provide better classification result.

The proposed method has been tested on a fully polarimetrie E-SAR (L-Band) data acquired on the. area of Penajam, East Kalimantan, Indonesia.

This paper is organized into the following fashions. Section II briefly describes the feature extraction procedures, which extracting several main polarimetrie parameters from polarimetrie covariance matrix and several physical scattering

characteristics based on Cloude's polarimetric decomposition. Section III explains the proposed classification procedure, which is designed by integrating spatial-contextual information with fuzzy maximum likelihood clustering algorithm. The experimental results are reported in section IV, and finally, Section V provides a discussion and conclusion.

2. Feature Extraction Schemes

2.1. Polarimetric Data Representation and Polarimetric Parameter Features Extraction

For radar polarimetry, the backscattering properties of the target can be completely described by a 2x2 complex scattering matrix, S, such that

$$S = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} \quad (1)$$

where S_{hv} is the scattering element of horizontal transmitting and horizontal receiving polarization, and the other three elements are similarly defined. For the reciprocal backscattering case, $S_{hv} = S_{vh}$. Because there are effectively only three independent elements, the polarimetric scattering information can also be represented by a target vector,

$$k = \begin{bmatrix} S_{hh} & \sqrt{2}S_{hv} & S_{vv} \end{bmatrix}^T \quad (2)$$

where the superscript "T" denotes the matrix transpose. The $\sqrt{2}$ on the S_{hv} term is to ensure consistency in the span (total power) computation. A polarimetric covariance matrix C can be formed by:

$$C = k.k^{*T} = \begin{bmatrix} |S_{hh}|^2 & \sqrt{2}S_{hh}S_{hv}^* & S_{hh}S_{vv}^* \\ \sqrt{2}S_{hv}S_{hh}^* & 2|S_{hv}|^2 & \sqrt{2}S_{hv}S_{vv}^* \\ S_{vv}S_{hh}^* & \sqrt{2}S_{vv}S_{hv}^* & |S_{vv}|^2 \end{bmatrix} \quad (3)$$

where the superscript "*" denotes the complex conjugate. C is a 3x3 Hermitian matrix, and has only six independent elements which can be employed as feature sets for classification purposes. Three real numbers on the main diagonal represent the powers (or intensity) of each polarization channels.

The other three complex numbers on the off diagonal represent the complex correlations, which can be used to quantify the similarity of waves (or coherence) at different polarization. For this purpose, the normalized value of this complex correlations (for example, for element C_{13} is given by $y = \langle S_{hh} S_{vv}^* \rangle$ /) is generally used. The magnitude of y (i.e., $|y|$) gives a measure of the degree of polarimetric coherence and lies between zero (incoherent) and one (completely coherent). The phase of y represents the phase difference between two polarization states and lies between 0 and 180°. The degree of polarimetric coherence and polarimetric phase difference closely related to the physical characteristics of the target scene so it can be used as a feature set to discriminate different land-cover types (Woodhouse, 2006). However, the sensitivity of these parameters are different depend on what polarizations are chosen. Hence, in this work, we select two most promising features: i.e., HH-VV polarimetric coherence and HH-VV polarimetric phase difference.

2.2. Feature Extraction based on Cloude's Polarimetric Decomposition

Fully polarimetric data provides unique possibility to separate scattering contributions of different nature, which can be associated to certain elementary scattering mechanisms (e.g. surface or single-bounce, double-bounce, and volume scattering). Several decomposition techniques have been proposed for this purpose. In recent years, approaches based

on the coherency matrix is usually preferred because its elements have relationship to the physics of wave scattering. The coherency matrix is formed, similarly with the covariance matrix C, but using the Pauli target vector k_p, such that

$$T = k_p . k_p^{*T} , \quad (4)$$

where

$$k_p = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{hh} + S_{vv} & S_{hh} - S_{vv} & 2S_{hv} \end{bmatrix}^T$$

The polarimetric decomposition theorem introduced by Cloude and Pottier (1997) proposed to identify polarimetric scattering mechanisms based on the eigenvalue analysis of a coherency matrix T. Applying eigenvalue analysis, the matrix T is decomposed into a sum of three coherence matrices T_i, each weighted by its corresponding eigenvalue λ_i.

$$T = \sum_{i=1}^{3} \lambda_i T_i = \lambda_1 (\mu_1 . \mu_1^{*T}) + \lambda_2 (\mu_2 . \mu_2^{*T}) + \lambda_3 (\mu_3 . \mu_3^{*T}) \quad (5)$$

Each matrix T_i is a unitary scattering matrix representing a deterministic scattering contribution. The amount of the contributions is given by the eigenvalues λ_i, while the type of scattering is related to the eigenvectors μ_i. The eigenvectors can be formulated as

$$\mu_i = \begin{bmatrix} \cos\alpha_i & \sin\alpha_i \cos\beta_i e^{j\delta_i} & \sin\alpha_i \sin\beta_i e^{j\gamma} \end{bmatrix}^T \quad (6)$$

The α angle corresponds to the continuous change from surface scattering ($\alpha = 0°$), moving into dipole or volume scattering ($\alpha = 45°$), moving into double bounce scattering between two dielectric surfaces, and finally reaching dihedral scatter from metallic surfaces at $\alpha = 90°$. The β angle is twice of the polarization orientation angle. The δ angle is the phase difference between the decomposed $S_{hh} + S_{vv}$ and $S_{hh} - S_{vv}$ terms, and the γ angle is the phase difference between the decomposed $S_{hh} + S_{vv}$ and S_{hv} terms. The ϕ angle is phase of the decomposed $S_{hh} + S_{vv}$ term.

Cloude and Pottier defined three secondary parameters, entropy H, anisotropy A, and mean alpha angle $\overline{\alpha}$, to characterize the result of the decomposition.

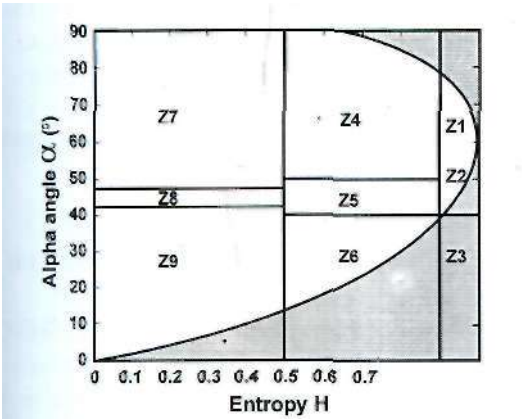
$$H = -\sum_{i=1}^{3} P_i \log_3 P_i \quad where \quad P_i = \frac{\lambda_i}{\sum_{j=1}^{3} \lambda_j} \quad (7)$$

$$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3} \quad (8)$$

$$\overline{\alpha} = \sum_{i=1}^{3} P_i \alpha_i \quad (9)$$

The entropy H, ranging from 0 to 1, represents the randomness of the scattering, with H = 0 indicating a single scattering mechanism (isotropic scattering) and H = 1 representing a random mixture of scattering mechanisms. For ocean and less rough surfaces, surface scattering will dominate, and H is near 0. For heavily vegetated areas, the H value will be high, due to multiple scattering mechanisms. The anisotropy A represents the relative importance of the second and third scattering mechanisms. A high anisotropy states that only the second scattering mechanism is important, while a low anisotropy indicates that the third scattering mechanism also plays a role. The mean alpha angle a reveals the averaged scattering mechanisms from surface scattering (a—»0°), volume scattering (a -> 45"), to double bounce scattering (a —> 90°). H and a clearly characterize the scattering characteristics of a medium. Cloude and Pottier further suggest an unsupervised classification scheme, using the H - a plane sub-divide

into 8 basic zones characteristic of different scattering behaviors, as shown in Figure 2. However, this unsupervised estimation of the type of scattering mechanisms may reach some limitations due to the arbitrarily fixed linear boundaries in the *H-a* plane which may not fit to data distribution, leading to noisy classification results (Ferro-Famil *et al.*, 2005; Lee *et al*, 1999-a). Hence, in this work, we use entropy *H*, anisotropy *A* , and mean alpha angle ***a*** directly as classification feature inputs to the fuzzy clustering classifier which will be described in Section 3.



Physical scattering characteristics:
29: Low Entropy Surface Scattering
28: Low Entropy Dipole Scattering
Z7: Low Entropy Multiple Scattering
26: Medium Entropy Surface Scattering
25: Medium Entropy Vegetation Scattering
Z4: Medium Entropy Multiple Scattering
23: (Not a Feasible Region in H- *OC* space)
22 : High Entropy Vegetation Scattering
21 : High Entropy Multiple Scattering

Figure 2. H - α pla

3. Proposed FMLE Clustering Including Spatial Context

3.1. Fuzzy Clustering

A fuzzy clustering algorithm yields a multiple-class pixel assignment where each pixel has membership in every class, but with varying degree. The memberships produce a fuzzy partition of the data that is viewed as an unsupervised classification. The following description is based on (Canty, 2006; Tso and Mather, 2001). The input features set (pixel vectors) $\{x_t,x_2,...,x_n\}$ consists of n vectors $x_t \in R^d$ (d is dimension of input features). Assuming there are K classes, $u_{kt} = u_k(x_t) \in [0,1]$ is the membership of the t th sample x_t in the k th class and $U = [u_{kt}]$ is the associated membership matrix. The set of cluster centers m_k is denoted by $m = (m_1, m_2, ..., m_K)$. Each sample point x_t satisfies the conservation of the membership constraint $u_{kt} \in [0,1]$ and

$$\sum_{k=1}^{K} u_{kt} = 1, \text{ for all } n \qquad (10)$$

The Fuzzy K-Means (FKM) clustering algorithm is based on minimization of the following fuzzy objective function

$$J_q(U,m) = \sum_{k=1}^{K}\sum_{t=1}^{n} u_{kt}^{q} D_{kt}^2 \qquad (11)$$

where D_{kt} is Euclidean distance $D^2_{kt}(x_t,m_k) = (x_t-m_k)$. The parameter q $(q>1)$ determines the "degrees of fuzziness" and is often chosen as $q = 2$. For $q > 1$ and $x_t \neq m_k$, for all i, k, a minimum of J_q may be achieved under the circumstance:

$$u_{ki} = \frac{\left(\frac{1}{|x_i - m_k|^2}\right)^{1/(q-1)}}{\sum_{k=1}^{K}\left(\frac{1}{|x_i - m_k|^2}\right)^{1/(q-1)}}, \text{ for all } i \qquad (12)$$

and the k th cluster mean is calculated from:

$$m_k = \frac{\sum_{i=1}^{n} u_{ki}^q x_i}{\sum_{i=1}^{n} u_{ki}^q}, \text{ for all } k \qquad (13)$$

The FKM clustering thus performed by iteratively applying Equations (12) and (13). The iteration terminates when the cluster centers m_k or alternatively when the matrix elements ju_{ki} cease to change significantly.

The FKM algorithm has been implemented successfully in many applications, such as pattern classification and image segmentation. However, the standard FKM algorithm is based on fuzzy objective function of Equation 11 that using Euclidean distance. This favor the formation of hyperspherical clusters having similar radii. An alternative algorithm, the FMLE algorithm (Canty, 2006) allows for ellipsoidal clusters of arbitrary extent and is consequently more flexible. The FMLE algorithm can be derived from FKM algorithm by replacing Equation (12) for the class membership $/.i_{ki}$ by the posterior probability $P(k|x_i)$ for class k given the observation x_i . That is, using Bayes' Theorem,

$$u_{ki} \rightarrow P(k \mid x_i) = \frac{P(x_i \mid k)P(k)}{P(x_i)} \qquad (14)$$

$P(x_i|k)$ is a class-specific probability density function, and is assumed to follow a multivariate normal distribution assumption. Its estimated mean m_k and covariance matrix s_k are given by:

$$m_k = \frac{\sum_{i=1}^{n} u_{ki} x_i}{\sum_{i=1}^{n} u_{ki}}$$

$$s_k = \frac{\sum_{i=1}^{n} u_{ki}(x_i - m_k)(x_i - m_k)^{\mathrm{T}}}{\sum_{i=1}^{n} u_{ki}}, \qquad (15)$$

$k = 1 \ldots K$

The prior probability of k th cluster $P(k)$, is given by:

$$P(k) = \frac{n_k}{n} = \frac{\sum_{i=1}^{n} \mu_{ki}}{n} \qquad (16)$$

Thus, apart from a constant factor independent of k, we obtain

$$u_{ki} = P(x_i|k)P(k) = \frac{1}{\sqrt{|s_k|}} \exp\left[-\frac{1}{2}(x_i - m_k)^{\mathrm{T}} s_k^{-1}(x_i - m_k)\right]\frac{n_k}{n} \qquad (17)$$

The FMLE algorithm consists of an iteration of Equations (10), (15), (16), and (17) until class membership converges. In this work, we use the following termination criterion:

$$\max\left(\left\|\mu_{ki}' - \mu_{ki}\right\|\right) < \varepsilon \qquad (18)$$

where ε is a threshold and often chosen as $\varepsilon = 0.001$.

Unlike the FKM algorithm, the memberships μ_{ki} are now functions of the directionally-sensitive Mahalanobis distance

$$D_M(x_i, m_k) = \sqrt{(x_i - m_k)^{\mathrm{T}} s_k^{-1}(x_i - m_k)}$$

Because of the exponential dependence on D_M^2 of the memberships in Equation (17), the computation is very sensitive to initialization conditions and can even become unstable. To avoid this problem,

we follow the suggestion of Gath and Geva (1989) and first obtain initial values for the $//_K$ by preceding the calculation with the FKM algorithm. For this purpose, in this work, we use only the entropy H and mean alpha angle $\boldsymbol{a}$ features to obtain initial fi_{tj}. However, the eight zones (or classes) in the H - $\boldsymbol{a}$ plane are not applied here, instead, we determine the number of classes based on the image content and a ground survey information.

3.2. **Spatial-contextual information based on Probabilistic Relaxation**

Both FKM and FMLE clustering algorithm described above make use exclusively the spectral (or intensity) properties of the individual pixel vectors and spatial-contextual information of the image was not taken into account.

In order to incorporate spatial-contextual information in classification process, in this paper, we adapted the probabilistic relaxation framework. This idea is based on the assumption that two neighbouring pixels are not entirely statistically independent: In reality, spatially random classification results are not very likely, instead continuous areas of certain sizes are to be expected. It seems clear that information from neighbouring pixels should increase the discrimination capabilities of the pixel-based measured data, and thus, improve the classification accuracy and the interpretation efficiency. Such ancillary information can be expressed by a neighbourhood function q, which must somehow reflect the contextual information of the neighbourhood (Canty, 2006). In order to define it, a compatibility coefficient

$$p\left(g,k_i \mid h,k_j\right) \tag{19}$$

is introduced, i.e., the conditional probability that pixel g falls into class k_i, if a neighbouring pixel h belongs into class k_j. As mentioned before, K possible class assignments are possible; furthermore it is possible to incorporate a larger neighbourhood consisting of L pixels. Based on this, a neighbourhood function

$$q(g,k_i) = \sum_{h=1}^{L}\sum_{j=1}^{K} p\left(g,k_i \mid h,k_j\right) p\left(h,k_j\right) \tag{20}$$

can be defined, which describes the total joint probability over all neighbours and their class assignments, that a pixel g falls into class k_l. The probability $q()$ gives information about class membership of pixel g solely by examination of its neighbourhood and without considering content of the pixel itself.

After the FMLE clustering procedure, the class membership probabilities (according to Equation 17) are known. This allows to evaluate Equation (20) and results in two kinds of class probabilities for each pixels: One, q_{kj}, based only on spatial contextual information, and another, fi_u, based on spectral information only. A combined spectral-spatial class membership for the next iteration of the FMLE is then determined by

$$\mu_{ki}^{'} = \frac{\mu_{ki} q_{ki}}{\sum_{k=1}^{K} \mu_{ki} q_{ki}} \tag{21}$$

Alternatively, this probabilistic relaxation process can be iteratively repeated before continue to the next iteration of the FMLE procedure in order to "propagate" (or to diffuse) the temporary updated results to their surrounding pixels / regions. In this work, we will observe the influence of number iterations of the probabilistic relaxation process on the classification result. An optimal number of iterations then will be determined by experiments.

The compatibility coefficients can be estimated from initially classified image. However, FMLE classifier is an iterative optimization classification procedure, then does not provide optimum result from first iteration. For this reason, in this proposed method, we apply spatial-contextual information after pixel-by-pixel basis FMLE classifier converges below a certain threshold £,. , and then continue using contextual FMLE classifier (FMLE following by probabilistic relaxation iteratively) until converges below a certain threshold £, . In this paper, we use

$S_x = 0.005$ and $s_2 = 0.001$.

4. Experimental Results

The proposed method is tested using single look complex (SLC) fully polarimetric-SAR data acquired over Penajam area, East Kalimantan Province. These data were acquired in L-band by Airborne E-SAR method on November 17th, 2004. **The** spatial resolution of the data used is 1.99 m and 3.0 m, in range and azimuth respectively. The scene under study contains different type of land covers: forest, fields, bare soils, and water area. Figure 3 shows a set of ground survey information. In Figure 3, the RGB image is formed using Pauli decomposition (Helmann, 2001).

For preprocessing, we construct scattering matrix from single look data (SLC) data for each polarization and then apply speckle filtering using J.S.Lee Polarimetric Filter (Lee *et **ai,*** 1999-b). In this experiment, a 3x3 window has been used. Larger windows provide more speckle smoothing but may smear fine details in the image.

Figure 4 shows the polarimetric parameters {5 features) extracted from covariance matrix. The powers of each polarization channels i.e.,: HH, HV, and VV intensity, are the most promising features for discriminating between different land-cover types. The HH-VV polarimetric coherence and HH-VV polarimetric phase difference have also such discrimination capability, but with relatively lower capability than intensity features. The HH-VV polarimetric coherence is particularly useful to discriminate forest (vegetation) area from other classes.

Figure 5 shows the features extraction results from Cloude's polarimetric decomposition. By analyzing mean alpha angle *ex* and entropy H, we can observe that open water area is characterized by surface scattering (alpha values less than 42.5°) with low entropy, while forest area is characterized by volume scattering (alpha values near 45°) with high entropy ($H > 0.9$). Bare soils and fields are both characterized relatively by medium entropy and low alpha values, and may cause low separability between these two classes. Anisotropy A does not provide sufficient sensitivity for the separation of the different land-cover types, however, may be used for separation of the bare soil class and field class.

We then use these two feature datasets (i.e., five features of the polarimetric parameter and three features of Cloude's polarimetric decomposition) and combined features of both datasets (totally eight features) as input for FMLE classifier. Two versions of the FMLE were applied in these experiments:

1. Non-contextual FMLE classification. In this case, probabilistic relaxation was not applied in all FMLE iterations. (The classification process is performed using only pixel-by-pixel basis classifier. See Figure 1.)
2. Contextual FMLE classification. In this case, 4-neighbourhood probabilistic relaxation was applied with 1, 3, 5, 7,

and 9 iterations. (The classification process is first performed using pixel-by-pixcl basis classifier, then continued using spatial-contextual classifier)

The classification results of the non-contextual FMLE systems are shown in Fig 6 (a, c, and e). We can observe that polarimctric parameter features alone can provide reasonable result, but with some misclassification between forest, fields, and bare soils. For example, the bare soil areas in upper left comers of the image (see Figure 6-a) were erroneously classified as forest. On die other hand, Cloude's decomposition features can identified accurately these bare soil areas (see Figure 6-c) and enhanced the discrimination between forest and non-forest areas. By combining these two feature datasets, the discrimination of different land cover types can be improved, thus giving belter classification result (see Figure 6-e).

As comparison, the classification results using standard FKM clustering (using Euclidean distance) are also presented in Figure 6 (b, d, and 0- in all results, we observed that the FMLE clustering perform consistently better than the FKM clustering. Some misclassification between forest, fields, and bare soils are occurred evidently. and water class at river areas can not be accurately identified by FKM clustering.

The classification results of the contextual FMLE systems are shown in Figure 7. In Figure 7-b, 1 iteration of the probabilistic relaxation have been used. Comparing with non-contextual result (Figure 7-a), although more homogeneous result is obtained, but the improvement is marginal. The classification results get more homogeneous (suppress more "salt and pepper" effect in homogeneous areas) by increasing the number of iteration. However, too many iterations lead to a widening of the effective neighbourhood of a pixel to such an extent that fully irrelevant spatial information falsifies the final classification results. It can also be confirmed in Figure 7-e (with 7 iterations) and Figure 7-f (with 9 iterations), which some erosion of the object boundaries (particularly when the objects are small in size) are occurred evidently. We conclude that the best results are obtained with 3-5 iterations (Figure 7-c or Figure 7-d). as it provide homogeneous classification result, but still preserve edge and other fine structures.

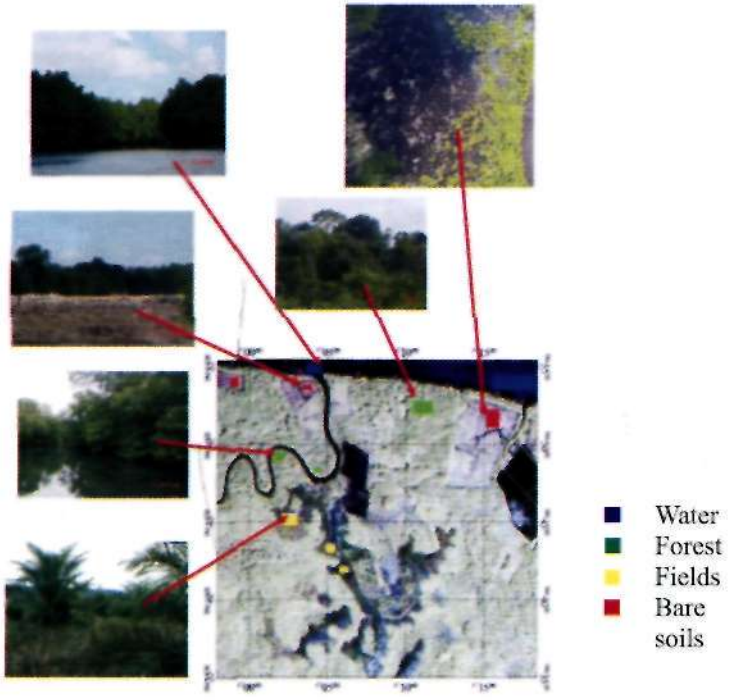
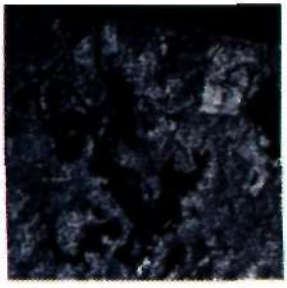
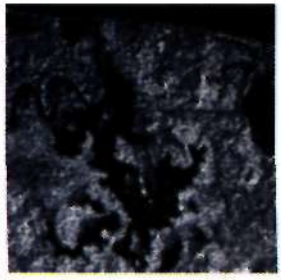


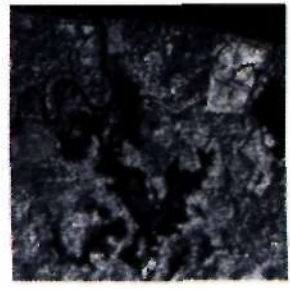
Figure 3. Ground Survey Information



a) HH Intensity (Polarisation power)

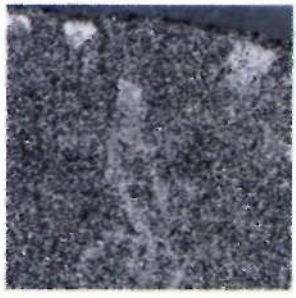


b) HV Intensity (Polarisation power)



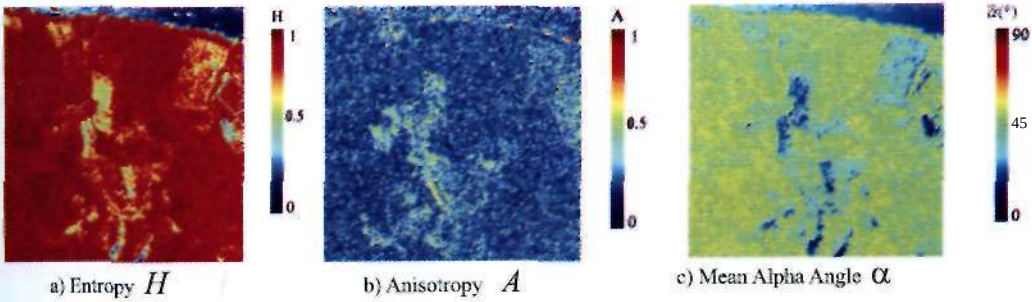
c) VV Intensity (Polarisation power)

d) HH-VV Polarimetric Coherence



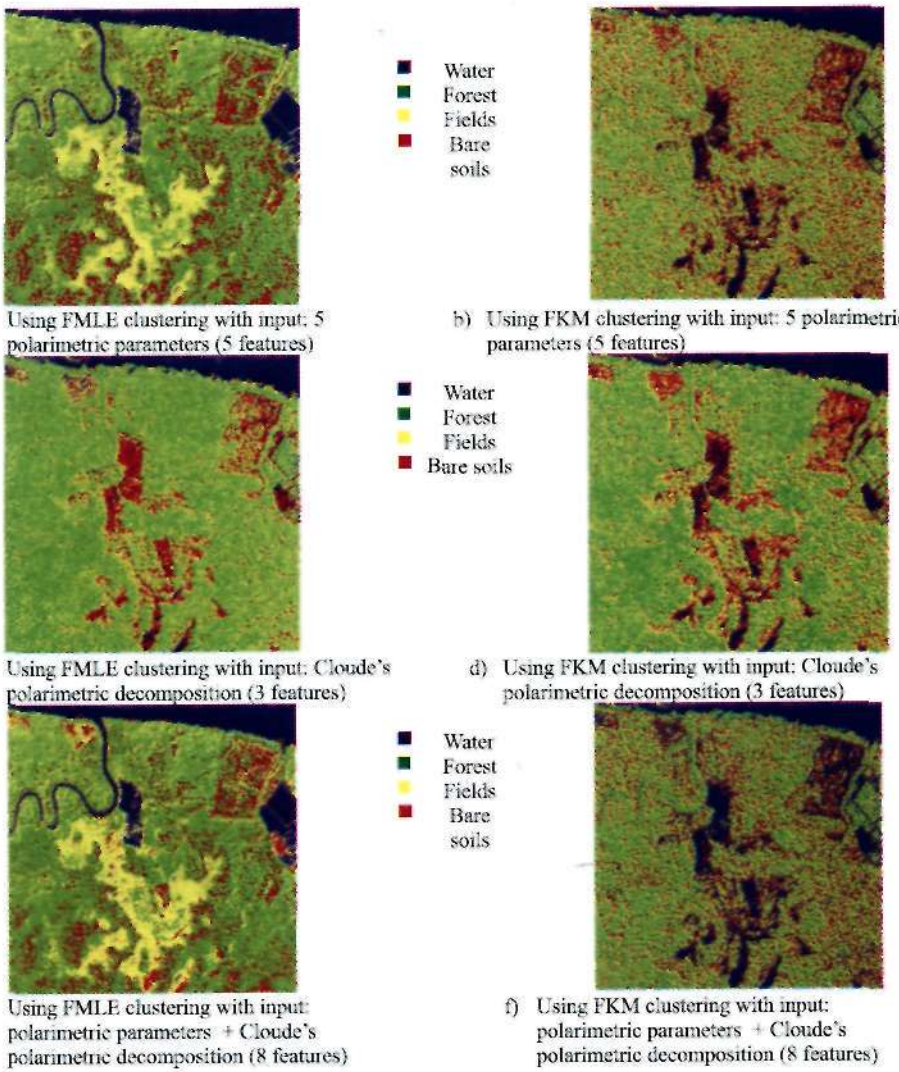
c) HH-VV Polarimetric Phase Difference

Figure 4. Polarimetric physical parameter features extracted from Polarimetric Covariance Matrix.



a) Entropy H b) Anisotropy A c) Mean Alpha Angle α

Figure 5. Features extracted from Cloude's decomposition model



a) Using FMLE clustering with input: 5 polarimetric parameters (5 features)

b) Using FKM clustering with input: 5 polarimetric parameters (5 features)

c) Using FMLE clustering with input: Cloude's polarimetric decomposition (3 features)

d) Using FKM clustering with input: Cloude's polarimetric decomposition (3 features)

e) Using FMLE clustering with input: polarimetric parameters + Cloude's polarimetric decomposition (8 features)

f) Using FKM clustering with input: polarimetric parameters + Cloude's polarimetric decomposition (8 features)

Figure 6. Classification results using non-contextual FMLE clustering with combined features of Cloude's polarimetric decomposition and polarimetric physical parameters. These results are obtained without spatial-contextual information. (Classification results using FKM clustering are also presented as comparison)

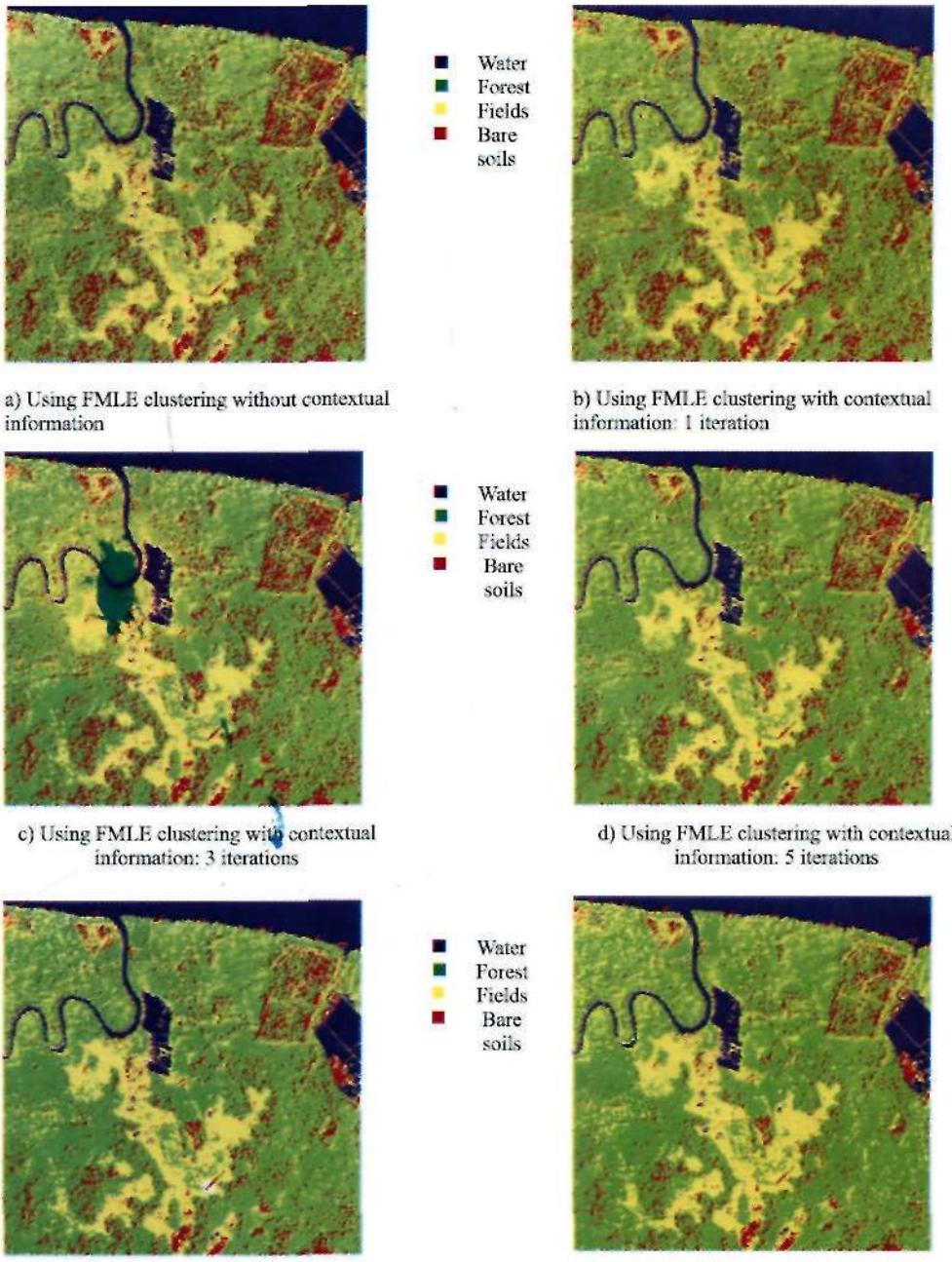


Figure 7. Classification results using FMLE clustering with combined features of Cloude's polarimetric decomposition (3 features) and polarimetric physical parameters (5 features).

5. Conclusion

An alternative method for unsupervised classification of polarimetric-SAR data has been proposed. The method was designed by integrating the combined features extracted from polarimetric covariance matrix and Cloude's polarimetric decomposition with contextual FMLE classifier.

The proposed method has been tested on a fully polarimetric, single look complex E-SAR (L-Band) data acquired on the area of Pcnajam, East Kalimantan, Indonesia. Experimental results show that the proposed method improves land-cover discrimination performance, and provides robust and homogeneous classification results but still preserving edge and other fine structures.

Acknowledgement

The authors would like to thank The Ministry of Forestry Republic of Indonesia for providing the E-SAR polarimetric data. The used polarimetric data set was acquired through IXDREX-II experiment (Indonesian Airborne Radar Experiment) supported by the European Space Agency.

References

Bruz^one, L., M. Marconcini, U. Wegmuller, and A. Wiesmann. 2004, An Advanced method for. the automatic classification of multitemporal SAR images, IEEE Trans. Geosci. Remote Sensing, 42(6): 1321-1334.

Canty, M. J., 2006, Image analysis, classification and change detection in remote sensing, CRC Press.

Cant>', M. J. and A. A. Nielsen, 2004, Unsupervised classification of changes in multispectral sattelite imagery, in Proceedings of SPIE, Image and Signal Processing for Remote Sensing X, 5573: 356-363.

Cloude. S. R. and E. Potticr, 2007. An entropy based classification scheme for land applications of polarimetric SAR, IEEE Trans. Geosci. Remote Sensing, 35(1): 68-78.

Davidson, G., K. Ouchi, G. Saito, N. Ishitsuka, N. Mohri, and S. Uratsuka, 2002, Polarimetric classification using expectation methods, Polarimetric and Interferometric SAR Workshop, Communications Research Laboratory, Tokyo.

Ferro-Famil, U E. Pottier, H. Skriver, P. Lumsdon, R. Moshammer, and K. Papathanassiou, 2005, Forest mapping and classification using L-Band Polinsar data in Proceedings of the Workshop on POLINSAR - Applications of SAR Polarimetry and Polarimetric Ihterfcromctry.

Gath.I.andA.B.Geva, 1989, Unsupervised optimal fuzzy clustering, IEEE Trans, on Pattern Analysis and Machine Intellegence, 11(7): 773-781.

Hellman, M. P., 2001, SAR Polarimetry Tutorial, Beta version 0.1a, SAR Polarimetry Tutorial.

Karathanassi, V. and M. Dabboor, 2004, Land cover classification using E-SAR polarimetric data in Proceedings of the XXthlSPRS Congress.

Kersten, P. R., J. S. Lee. and T. L. Ainsworth, 2005, Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering, IEEE Trans. Geosci. Remote Sensing, 43(3): 519-527.

Lee, J. S.,M.R.Grunes,T.L. Ainsworth, L.J. Du, D. L.__Schuler, and E. R. Cloude, 1999-a. Unsupervised classification using polarimetric decomposition and the complex Wishart classifier, IEEE Trans. Geosci. Remote Sensing, 37(5): 2249-2258.

Lee, J. S., M. R. Grunes, and G. D. Grandi, 1999-b, Polarimetric SAR speckle filtering and its impact on terrain classification, IEEE Trans. Geosci. Remote Sensing, 37(5): 2363-2373.

Lumsdon, P., 2003, Land cover classification and height estimation in polarimetric SAR interferometry, in Proceedings of the Workshop on POLINSAR - Applications of SAR Polarimetry and Polarimetric Interferometry.

Sambodo, K. A., A. Murni, and M. Kartasasmita, 2007, Classification of polarimetric-SAR data with neural network using combined features extracted from scattering models and texture analysis, International Journal of Remote Sensing and Earth Sciences, 4:1-17.

Scheuchl. B., R. Caves, I. dimming, and G. Staples, 2001, H/Aa-based classification of sea ice using SAR polarimetry, in Proceedings of the 23rd Canadian Symposium on Remote Sensing.

Tso, B., and P. M. Mather. 2001, Classification Methods for Remotely Sensed Data,Taylor& Francis Inc.

Woodhouse, I., 2006, Introduction to Microwave Remote Sensing, CRC Press.