

DUAL FREQUENCY POLARIMETRIC SAR DATA CLASSIFICATION AND ANALYSIS

L. Ferro-Famil

Ecole Polytechnique de l'Université de Nantes
IRESTE, Laboratoire SEI, EP CNRS 2018
Rue C. Pauc, BP 50609, 44306 Nantes Cedex, France

E. Pottier

Université de Rennes 1
UPRES-A CNRS 6075 “Structures Rayonnantes”
Laboratoire Antennes et Télécommunications
Campus de Beaulieu, Bât 22, 263 Av Général Leclerc
CS74205, 35042 Rennes cedex, France

Abstract—In this paper, we introduce a new classification scheme for dual frequency polarimetric SAR data sets. A (6×6) polarimetric coherency matrix is defined to simultaneously take into account the full polarimetric information from both images. This matrix is composed of the two coherency matrices and their cross-correlation. A decomposition theorem is applied to both images to obtain 64 initial clusters based on their scattering characteristics. The data sets are then classified by an iterative algorithm based on a complex Wishart density function of the 6 by 6 matrix. A class number reduction technique is then applied on the 64 resulting clusters to improve the efficiency of the interpretation and representation of each class characteristics. An alternative technique is also proposed which introduces the polarimetric cross-correlation information to refine the results of classification to a small number of clusters using the conditional probability of the cross-correlation matrix. The analysis of the resulting clusters is realized by determining the rigorous change in polarimetric properties from one image to the other. The polarimetric variations are parameterized by 8 real coefficients derived from the decomposition of a special unitary operator on the Gell-Mann basis. These classification and analysis schemes are applied to full polarimetric P , L , and C bands SAR images of the Nezer forest acquired by NASA/JPL AIRSAR sensor (1989).

- 1. Introduction**
 - 2. Dual Frequency Polarimetric Sar Data Statistics**
 - 2.1 Dual Polarimetric Representation
 - 2.2 Cross-Correlation Matrix Statistics
 - 3. Single Image Classification Procedure**
 - 3.1 Classification Algorithm
 - 3.2 Application to POLSAR Data
 - 4. Multi Frequency Classification Procedure**
 - 4.1 Dual Image Maximum Likelihood Classification
 - 4.1.1 *Classification Algorithm*
 - 4.1.2 *Reduction of the Number of Classes*
 - 4.1.3 *Application to POLSAR Data*
 - 4.2 Dual Image Classification Using the Cross-Correlation Information
 - 4.2.1 *Classification Algorithm*
 - 4.2.2 *Application to POLSAR Data*
 - 5. Dual Polarimetric Class Description**
 - 5.1 Special Unitary Transformation
 - 5.2 Application to POLSAR Data
 - 6. Conclusion**
- References**

1. INTRODUCTION

The backscattering properties of a natural medium vary with the observation frequency according to its physical features, such as its structure or its dimensions with respect to the radar wavelength. The scattering mechanism may remain almost unchanged for bare soil observation at L and C bands, but may show a totally different aspect for forest remote sensing at P and C bands. Incident waves with different wavelength interact with separate parts of a complex medium. The purpose of multi-frequency analysis is to gather adequate information from each data set.

Many algorithms have been developed to classify natural media using polarimetric synthetic aperture radar (POLSAR) data [1–5]. Several approaches were derived to directly relate some basic characteristics of the targets to elements of the polarimetric covariance matrix [6–8]. More recently, polarimetric decomposition theorems were introduced in order to investigate the intrinsic physical properties of a natural medium by evaluating the underlying scattering mechanisms

[9–12]. All these approaches realize an interpretation of the polarization of the backscattered wave, and establish a relation between the medium physical properties and polarimetric transformations.

The use of multi-frequency polarimetric data sets has been shown to increase the interpretation capabilities of quantitative remote sensing of natural media [13, 14]. Some multi-frequency full polarimetric classification approaches were developed using various types of algorithms and techniques based on neural networks, fuzzy iterative classifier, statistical segmentation, etc., [15–17]. Statistical classification using multivariate probability density functions permits us to define adaptive decision rules to segment data sets into more compact clusters in an unsupervised way. Moreover, a decision rule derived from a full polarimetric representation leads to optimal results and provides information for class type identification by evaluating the underlying physical scattering mechanism [17].

Kong et al. [18] introduced a maximum likelihood decision rule based on the multivariate complex gaussian distribution of the elements of the coherent scattering matrix. In order to limit the effects of speckle in polarimetric SAR images, data are generally processed through incoherent averaging. The polarimetric information of the averaged targets is represented by coherency matrices. Lee et al. [17] introduced the maximum likelihood decision rule to the incoherent case by using the multivariate complex Wishart distribution of sample coherency matrices. A k-mean algorithm was applied to iteratively assign the pixels of the POLSAR image to one of the different classes using the maximum likelihood rule. Lee et al. [19] further improve the classification by using the H - α decomposition theorem [9] to provide an initial guess of the pixel distribution into the classes that produces a better convergence of the unsupervised classification algorithm.

In this paper, we propose an unsupervised classification of dual frequency POLSAR images by including the polarimetric information of both images. A (6×6) coherency matrix is constructed using the single look complex data from the two frequency images. This matrix includes the coherency matrices from each image as well as their cross-correlation [20, 21]. This matrix is shown to follow a Wishart distribution and a maximum likelihood decision rule is derived. Similarly to the single image case, data sets are processed through a k-mean classifier after an initialization step consisting in the application of the H - α classification procedure to each separate image. A class number

reduction technique is applied on the 64 resulting clusters to improve the efficiency of the interpretation and representation of each class characteristics. An alternative approach is also proposed based on the introduction of the second image polarimetric information through the cross-correlation conditional statistics. This procedure permits an efficient handling of the classification information by refining an initial classified data set with a small number of clusters and iteratively creating new classes. The probability density of a sample cross-correlation matrix conditionally to the polarimetric information in one image is derived in order to calculate a distance measure. This distance is used to perform an unsupervised splitting of a dual image cluster into 2 sub-sets. The resulting dual classes are analyzed by determining the variation of the scattering mechanism polarimetric properties from one data set to the other. A target vector transforms to another one by the way of a special unitary operator. This operator is parameterized in terms of 8 real coefficients obtained from a decomposition onto the Gell-Mann basis matrix set. The determination of the variation coefficients necessitate to resolve an over-determined set of non-linear equations.

The classification and analysis schemes are applied to full polarimetric P , L , and C bands SAR images of the Nezer forest acquired by NASA /JPL AirSAR sensor (1989).

2. DUAL FREQUENCY POLARIMETRIC SAR DATA STATISTICS

When dealing with dual frequency images, the polarimetric information contained in a resolution cell represents the fully polarimetric characteristics of both data sets.

2.1 Dual Polarimetric Representation

For a given measurement configuration, a target is fully characterized by its coherent target vector, $\mathbf{k}$, obtained using a straightforward lexicographic ordering of the q scattering matrix elements.

$$\mathbf{k} = [S_{HH} \ S_{HV} \ S_{VH} \ S_{VV}]^T \quad (1)$$

In the case of dual polarimetric data classification, it is important to simultaneously take into account the polarimetric information from

both images :

$$\mathbf{w} = \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{k}_2 \end{bmatrix} \quad (2)$$

where $\mathbf{k}_1$ and $\mathbf{k}_2$ are the target vectors belonging to the different images. The vector $\mathbf{w}$ has the dimension $\mathbf{p} = 2\mathbf{q}$. The $(p \times p)$ n -look covariance matrix $\mathbf{A}$ summarizes the joint information from both images and has the following structure:

$$\mathbf{A} = \frac{1}{n} \sum_{j=1}^n \mathbf{w}_j \mathbf{w}_j^\dagger = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad \text{with} \quad \mathbf{A}_{rs} = \frac{1}{n} \sum_{j=1}^n \mathbf{k}_{rj} \mathbf{k}_{sj}^\dagger \quad (3)$$

The matrices $\mathbf{A}_{11} = \mathbf{Z}_1$ and $\mathbf{A}_{22} = \mathbf{Z}_2$ are the standard n -look $(q \times q)$ covariance matrices from separate images. $\mathbf{A}_{12} (= \mathbf{A}_{21}^\dagger)$ is a $(q \times q)$ complex matrix containing information about the polarimetric cross-correlation between $\mathbf{k}_1$ and $\mathbf{k}_2$. The target vector $\mathbf{w}$ follows a complex normal distribution $N_C(\mathbf{0}, \boldsymbol{\Sigma}_w)$ [22], with $\boldsymbol{\Sigma}_w = E(\mathbf{w} \mathbf{w}^\dagger)$ its $(q \times q)$ covariance matrix. The sample $(p \times p)$ covariance matrix $\mathbf{A}$ has a complex Wishart distribution $W_C(n, \boldsymbol{\Sigma}_A)$, characterized by n degrees of freedom and by its covariance matrix $\boldsymbol{\Sigma}_A = \boldsymbol{\Sigma}_w/n$.

$$p(\mathbf{A}) = \frac{|\mathbf{A}|^{n-p} \exp(-\text{tr}(\boldsymbol{\Sigma}_A^{-1} \mathbf{A}))}{K(n, p) |\boldsymbol{\Sigma}_A|^n} \quad \text{with} \quad K(n, p) = \pi^{p(p-1)/2} \prod_{i=1}^p \Gamma(n-i+1) \quad (4)$$

The advantage in using the $(p \times p)$ representation resides in the fact that according to (4), dual data sets can be simultaneously classified by using the maximum likelihood distance measure defined in (5) and without any assumption concerning their independence. A pixel is assigned to the class X_m if $d_1(\mathbf{Z}, X_m) \leq d_1(\mathbf{Z}, X_j) \forall j \neq m$, with

$$d_1(\mathbf{A}, X_m) = \ln |\boldsymbol{\Sigma}_{Am}| + \text{tr}(\boldsymbol{\Sigma}_{Am}^{-1} \mathbf{A}) \quad (5)$$

with $\boldsymbol{\Sigma}_{Am}$ the $(p \times p)$ feature matrix of class X_m .

2.2 Cross-Correlation Matrix Statistics

The cross-correlation is highly sensitive to the scattering phenomenon type and its degree of randomness which both can be extracted from this covariance matrix representation. The probability of $\mathbf{A}_{12}$ may then be taken into account to separate groups of pixels belonging

to the same class but possessing slightly different dual polarimetric information.

The properties of the hermitian matrix $\mathbf{A}$ permit us to define a conditional probability of the polarimetric cross-correlation matrix $\mathbf{A}_{12}$. In order to express the conditional probability of $\mathbf{A}_{12}$ given $\mathbf{A}_{22}$, $p(\mathbf{A}_{12}|\mathbf{A}_{22})$, both $\mathbf{A}$ and $\Sigma_{\mathbf{A}}$ can be partitioned as follows [23]

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} \mathbf{A}_{11.2} + \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \\ \Sigma_{\mathbf{A}} &= \begin{bmatrix} \Sigma_{\mathbf{A}_{11.2}} + \Sigma_{\mathbf{A}_{12}}\Sigma_{\mathbf{A}_{22}}^{-1}\Sigma_{\mathbf{A}_{21}} & \Sigma_{\mathbf{A}_{12}} \\ \Sigma_{\mathbf{A}_{21}} & \Sigma_{\mathbf{A}_{22}} \end{bmatrix}\end{aligned}\quad (6)$$

with $\mathbf{A}_{11.2} = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$ and $\Sigma_{\mathbf{A}_{11.2}} = \Sigma_{\mathbf{A}_{11}} - \Sigma_{\mathbf{A}_{12}}\Sigma_{\mathbf{A}_{22}}^{-1}\Sigma_{\mathbf{A}_{21}}$. By the way of an upper-triangular transformation, the determinants of the $(p \times p)$ matrices are then simplified to the following expressions:

$$|\mathbf{A}| = |\mathbf{A}_{22}||\mathbf{A}_{11.2}| \quad \text{and} \quad |\Sigma_{\mathbf{A}}| = |\Sigma_{\mathbf{A}_{22}}||\Sigma_{\mathbf{A}_{11.2}}| \quad (7)$$

After some reductions and combinations, an expression of $\text{tr}(\Sigma_{\mathbf{A}}^{-1}\mathbf{A})$ is found as a function of $\mathbf{A}_{11.2}$, $\mathbf{A}_{12}$ and $\mathbf{A}_{22}$, with:

$$\begin{aligned}\text{tr}(\Sigma_{\mathbf{A}}^{-1}\mathbf{A}) &= \text{tr}\left[\Sigma_{\mathbf{A}_{11.2}}^{-1}(\mathbf{A}_{12} - \Sigma_{\mathbf{A}_{12}}\Sigma_{\mathbf{A}_{22}}^{-1}\mathbf{A}_{22})\right. \\ &\quad \cdot \mathbf{A}_{22}^{-1}(\mathbf{A}_{12} - \Sigma_{\mathbf{A}_{12}}\Sigma_{\mathbf{A}_{22}}^{-1}\mathbf{A}_{22})^H \Big] \\ &\quad + \text{tr}[\Sigma_{\mathbf{A}_{22}}^{-1}\mathbf{A}_{22}] + \text{tr}[\Sigma_{\mathbf{A}_{11.2}}^{-1}\mathbf{A}_{11.2}]\end{aligned}\quad (8)$$

Inserting (7) and (8) into the Wishart probability function of $\mathbf{A}$ in (4), it can be shown that:

$$p(\mathbf{A}) = p(\mathbf{A}_{11.2}, \mathbf{A}_{12}, \mathbf{A}_{22}) = p(\mathbf{A}_{11.2})p(\mathbf{A}_{12}, \mathbf{A}_{22}) \quad (9)$$

with

$$p(\mathbf{A}_{11.2}) = \frac{|\mathbf{A}_{11.2}|^{(n-q-q)}}{K(n-q, q)|\Sigma_{\mathbf{A}_{11.2}}|^{(n-q)}} \exp[-\text{tr}(\Sigma_{\mathbf{A}_{11.2}}^{-1}\mathbf{A}_{11.2})] \quad (10)$$

and

$$\begin{aligned}& p(\mathbf{A}_{12}, \mathbf{A}_{22}) \\ &= \frac{|\mathbf{A}_{22}|^{(n-q)}}{K(n, q)|\Sigma_{\mathbf{A}_{22}}|^n} \exp[-\text{tr}(\Sigma_{\mathbf{A}_{22}}^{-1}\mathbf{A}_{22})] \\ &\quad \cdot \frac{\exp[-\text{tr}(\Sigma_{\mathbf{A}_{11.2}}^{-1}(\mathbf{A}_{12} - \Sigma_{\mathbf{A}_{12}}\Sigma_{\mathbf{A}_{22}}^{-1}\mathbf{A}_{22})\mathbf{A}_{22}^{-1}(\mathbf{A}_{12} - \Sigma_{\mathbf{A}_{12}}\Sigma_{\mathbf{A}_{22}}^{-1}\mathbf{A}_{22})^H)]}{\pi^{q^2}|\Sigma_{\mathbf{A}_{11.2}}|^q|\mathbf{A}_{22}|^q}\end{aligned}\quad (11)$$

From (9) and (10), it is found that the matrix $\mathbf{A}_{11.2}$ is independent of $\mathbf{A}_{12}$ and $\mathbf{A}_{22}$, and follows a complex Wishart probability density function $W_C(n - q, \Sigma_{\mathbf{A}_{11.2}})$.

As the sample covariance matrix of the second image, $\mathbf{A}_{22}$, follows a complex Wishart density function with n degrees of freedom $W_C(n, \Sigma_{\mathbf{A}_{22}})$, and since $p(\mathbf{A}_{12}, \mathbf{A}_{22}) = p(\mathbf{A}_{12}|\mathbf{A}_{22})p(\mathbf{A}_{22})$, the conditional probability density of the polarimetric cross-correlation matrix $\mathbf{A}_{12}$ given $\mathbf{A}_{22}$, is a complex normal function, given by the last line of (11). This complex normal density function may be formulated as [23]:

$$p(\mathbf{A}_{12}|\mathbf{A}_{22}) = N_C(\mathbf{M}_{\mathbf{A}_{12}|\mathbf{A}_{22}}, \Sigma_{\mathbf{A}_{12}|\mathbf{A}_{22}}), \quad \text{with} \quad (12)$$

$$\mathbf{M}_{\mathbf{A}_{12}|\mathbf{A}_{22}} = \Sigma_{\mathbf{A}_{12}} \Sigma_{\mathbf{A}_{22}}^{-1} \mathbf{A}_{22} \quad \text{and} \quad \Sigma_{\mathbf{A}_{12}|\mathbf{A}_{22}} = \Sigma_{\mathbf{A}_{11.2}}^T \otimes \mathbf{A}_{22}$$

Equation (12) can be written under a conventional form using the properties of the Kroneker product $\otimes$

$$p(\mathbf{A}_{12}|\mathbf{A}_{22}) = p(\mathbf{y}|\mathbf{A}_{22}) = \frac{\exp\left[-(\mathbf{y} - \mathbf{m})^\dagger (\Sigma_{\mathbf{A}_{11.2}}^T \otimes \mathbf{A}_{22})^{-1} (\mathbf{y} - \mathbf{m})\right]}{\pi^{q^2} |\Sigma_{\mathbf{A}_{11.2}}^T|^q |\mathbf{A}_{22}|^q} \quad (13)$$

with $\mathbf{y} = \text{vec}(\mathbf{A}_{12}^\dagger)$ a q^2 complex element vector obtained by stacking the columns of $\mathbf{A}_{12}^\dagger$ under each other, and $\mathbf{m} = \text{vec}(\mathbf{M}_{\mathbf{A}_{12}|\mathbf{A}_{22}}^\dagger)$. This complex normal density function gives, for pixels belonging to the class X_m , the probability of the cross-correlation matrix conditionally to the second image sample covariance matrix. The logarithm of (13) is used to define a measure of the distance between the actual polarimetric cross-correlation $\mathbf{A}_{12}$ and its expected value when observing the second image polarimetric information $\mathbf{A}_{22}$, with :

$$d_2(\mathbf{A}_{12}|\mathbf{A}_{22}, X_m) = \text{tr}(\Sigma_{\mathbf{A}_{11.2}}^{-1}(\mathbf{A}_{12} - \mathbf{M}_{\mathbf{A}_{12}|\mathbf{A}_{22}}) \cdot \mathbf{A}_{22}^{-1}(\mathbf{A}_{12} - \mathbf{M}_{\mathbf{A}_{12}|\mathbf{A}_{22}})^\dagger) + q \ln(|\mathbf{A}_{22}|) \quad (14)$$

3. SINGLE IMAGE CLASSIFICATION PROCEDURE

3.1 Classification Algorithm

For a reciprocal medium in a monostatic radar configuration, the target vector presented in (1) is modified as $\mathbf{k}_p = \frac{1}{\sqrt{2}}[S_{HH} + S_{VV}$

$S_{HH} - S_{VV} - 2S_{HV}]^T$ leading to the definition the coherency matrix $\mathbf{T} = E(\mathbf{k}_p \mathbf{k}_p^\dagger)$.

The use of the coherency matrix instead of the covariance matrix does not modify the density distribution types and the related equations defined previously [19].

The method used to perform a classification of a single image polarimetric data set is based on the use of an iterative k -mean algorithm and is described in details in [17, 19].

Four options have to be chosen by the user:

- The number of classes: m .
- The initialization of the pixel distribution into the m classes.
- The distance measure from a pixel to the m class centers.
- The termination criterion.

The number of classes and the data assignment during the initialization step of the classification are critical points and determine the quality of the whole classification.

In [19], Lee et al. proposed to initialize the classes using the $H-\alpha$ classification scheme [9, 10], which provides 8 classes relating to the underlying physical scattering mechanism. This splitting of the data set gives a stable initial approximation. The distance is estimated using the maximum likelihood approach applied on the data statistics mentioned above. The termination criterion may be selected from the estimation of the classification quality, when a maximum number of iterations is reached, or when a sufficiently low number of pixels are switching classes from one iteration to the next.

3.2 Application to POLSAR Data

The classification technique is applied on the Nezer site situated in the Landes Forest in the south west of France. On August 16, 1989, full polarimetric data have been acquired by Nasa JPL AIRSAR sensor at P , L , and C bands, with frequency at 0.44 GHz, 1.225 GHz and 5.3 GHz, respectively. The pixel spacings are 3 m by 6.6 m.

The scene contains bare soil areas and many homogeneous forested areas of maritime pines. Several tree-age groups are included from more than 41 years down to 5–8 years of age. Backscattering from the tree parcels is highly correlated to the age of the trees.

Figure 2 shows the span images of the Nezer site at P band.

Figure 3 shows the results of the unsupervised Wishart classification

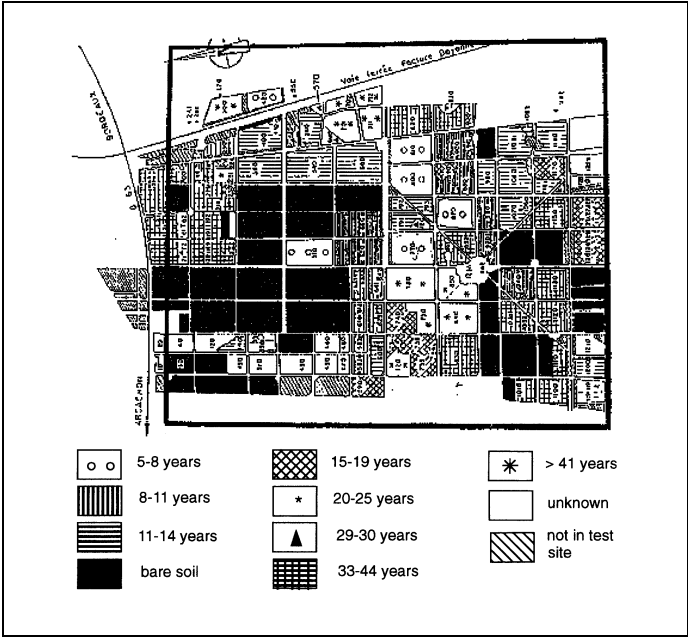


Figure 1. Ground truth of Nezer Forest. This map is a courtesy of CESBIO and Dr. Thuy Le Toan.



Figure 2. Span image at P band.



Figure 3. Results of the Wishart classification at P band (top) and L band (bottom).

applied on P band and L band POLSAR data sets.

After 4 iterations, the results obtained using the unsupervised Wishart classifier depict an important improvement in the classification accuracy. The P band classified image shows an interesting correspondence between the polarimetric classes and the ground truth information. Classes 7 and 8 represent the clear cut areas whereas cells covered with 5–8 years old trees are occupied by class 6. Medium age trees, from 11 to 19 years old, are mainly covered by the polarimetric classes 4, 5 and 6. Older tree cells, from 20 to more than 41 years old, have a polarimetric behavior corresponding to classes 1, 2, 3. A polarimetric class may spread over more than one type of forest cell tree.

The L band classified image does not separate different tree classes accurately. Young trees mainly correspond to class 5, while the rest of trees are represented by classes 1, 2, 3, and 4. The unsupervised classification at this frequency highlights differences within the clear-cut regions which are segmented in two different classes, 7 and 8.

This unsupervised classification algorithm modifies the decision boundaries in an adaptive way to better fit the natural distribution of the scattering mechanisms and takes into account the whole polarimetric information contained in the coherency matrix representation.

The characterization and interpretation of the different clusters may be achieved by studying the polarimetric properties of their center feature matrix from the parameters delivered by the $H-A-\alpha$ decomposition or other full polarimetric analysis techniques [19, 24].

4. MULTI FREQUENCY CLASSIFICATION PROCEDURE

We propose different approaches to the classification of several polarimetric data sets simultaneously. Dual polarimetric images are classified by the way of clustering procedures from the distance measure defined in (5) using the $(p \times p)$ coherency matrix representation, or successively (5) and (14) if the dual polarimetric information is introduced through the conditional probability density of the cross-correlation matrix.

4.1 Dual Image Maximum Likelihood Classification

4.1.1 Classification Algorithm

Both separate images are classified through the unsupervised Wishart classifier into 8 classes each. The results are further segmented into 64 classes by simultaneously considering the labels of pixels in each image. This way of initializing the pixel distribution presents the advantage to give equal significance to the polarimetric information interpretation from each image.

The initial classified image, made of 64 clusters, is then processed through an unsupervised k -mean clustering algorithm based on the distance measure defined in (5). The use of the (6×6) dual polarimetric coherency matrix permits to calculate in an easy way the distance from a pixel to the different class center feature matrices.

The dual image classification algorithm corresponding to the synopsis described in figure 4 is:

- Step 1: Perform the 8 class unsupervised Wishart classification on both separate polarimetric data sets.
- Step 2: Initialize the class distribution by calculating the combined class number using the following rule: A pixel belonging to

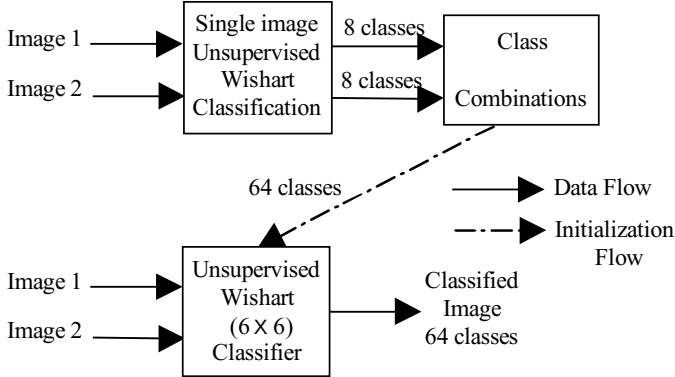


Figure 4. Synopsis of the dual polarimetric SAR data classification procedure.

class X_{1i} in the first image and to X_{2j} in the second one is assigned to the combined class X_{i+8j} . The number of classes is then equal to 64.

Step 3: For each class X_i , compute the 6 by 6 feature covariance matrix $\Sigma_i = \frac{1}{N_i} \sum_{N_i} \langle \mathbf{A} \rangle \in X_i$.

Step 4: Assign each pixel to the class minimizing the distance measure $d_1(\langle \mathbf{A} \rangle, X_m)$ from (5) over the 64 classes.

Step 5: Stop if a termination criterion is met, otherwise go to Step 3.

The accuracy, using this initialization, is highly enhanced, since pixels are distributed into 64 classes according to the combination of the full polarimetric unsupervised Wishart classification results obtained for separate images.

4.1.2 Reduction of the Number of Classes

The number of classes has to be reduced in order to facilitate the interpretation for each class characteristics as well as the visual representation of the geographical location of the different clusters. Lee et al. [19] proposed a merging procedure based on the study of the class compactness and separability that is well adapted to the natural partition of the data. Considering the whole class set, the clusters to be merged are the ones presenting the lowest degree of separability. Two classes can be distinguished if they are compact and if the mean distance between their elements is high, hence the separability between classes X_i and X_j is defined as the ratio of their between-class dis-

tance to their within-class dispersion [19].

For a given class X_i , the within-class dispersion W_i is defined as the mean distance from its elements to the (6×6) class center feature matrix Σ_i

$$W_i = \frac{1}{N_i} \sum_{N_i} d_1 [(\langle \mathbf{A} \rangle \in X_i), X_i] \quad (15)$$

Inserting the definition of the distance measure of (5) in (15), one may find

$$W_i = \ln|\Sigma_i| + \text{tr}(\mathbf{I}) = \ln|\Sigma_i| + z \quad (16)$$

The constant term ($z = 6$) corresponds to the trace of the (6×6) identity matrix $\mathbf{I}$.

The distance between class X_i and class X_j is the mean distance from the elements of each class to the center feature matrix of the other class.

$$B_{ij} = \frac{\frac{1}{N_i} \sum_{N_i} d_1 [(\langle \mathbf{A} \rangle \in X_i), X_j] + \frac{1}{N_j} \sum_{N_j} d_1 [(\langle \mathbf{A} \rangle \in X_j), X_i]}{2} \quad (17)$$

Using the definition of the distance measure, the average distance between two classes is simplified to

$$B_{ij} = \frac{W_i + W_j + \text{tr}(\Sigma_i^{-1} \Sigma_j + \Sigma_j^{-1} \Sigma_i)}{2} \quad (18)$$

Their separability $Sp(X_i, X_j)$ is then given by

$$Sp(X_i, X_j) = B_{ij}/(W_i + W_j) \quad (19)$$

Using this definition, and considering that the classes to be merged are these presenting the lowest separability, the class reduction technique is applied by the way of an iterative algorithm till a termination criterion is met.

One may use the termination introduced in [19] based on an estimation of the classification quality, or consider that the reduction procedure may end when an arbitrarily fixed number of classes is reached, so that the classification results can be efficiently handled.

4.1.3 Application to POLSAR Data

The classification algorithm is run on P , L and C , band data sets combinations, with a number of classes reduced to 16. The classification results of P band and L band data sets are shown in figure 5. In

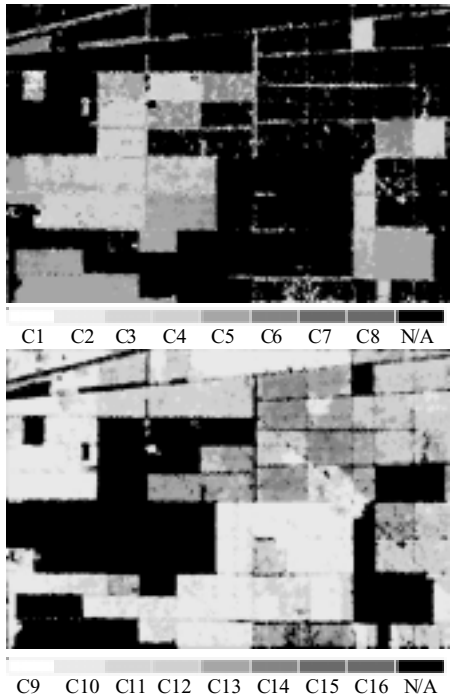


Figure 5. Results of the unsupervised P and L band dual data classification.

order to facilitate the representation of a 16 classes image in grayscale, the results are split into 2 images containing each 8 classes.

When comparing this classification method results with the ones obtained with a single image classification procedure, one finds an important improvement in the description of the natural characteristics distribution. The initialization step, which simultaneously takes into account both data sets with equal importance, produces a good discrimination of details like ways and small forest cells. The class number reduction technique permits us to merge classes of close characteristics, and produce accurate distribution of the differently aged trees in the forest.

The classification map obtained from P and L band images , shows a good concordance with the parcel distribution given by the ground truth information in Figure 1. The relevant information from each data set has been gathered to differentiate the different parcel types and to perform a discrimination between the different types of ways between

the parcels. Globally, the P band features permit us to discriminate the different forest textures, while L or C bands characteristics emphasize the roads and ways as well as the different types of bare ground areas. The compactness of the different classes is highly enhanced by this dual classification procedure.

It is important to note that no assumptions were made concerning the level of cross-correlation between the different data sets, the classification taking into account this information during the whole process.

4.2 Dual Image Classification Using the Cross-Correlation Information

4.2.1 Classification Algorithm

The procedure described above classifies data by reducing a high number of classes obtained by the combination of class labels in each image. The quality of classification depends on the accuracy of the separability criterion mentioned in (19). This top down approach can be computationally intensive since it manipulates up to 64 classes.

A problem linked to the reduction of the number of classes may be encountered when classifying scenes composed of various types of scatterers. In the case of forest remote sensing, point targets or classes corresponding to heterogeneities may be considered as highly separable clusters, while the response of the different types of forest parcels may appear to be very close. During the class number reduction process, the forest parcels may then be merged into classes containing a large number of pixels.

We propose another approach, which instead of initializing the classification with a high number of classes, begins with a small amount of classes and iteratively uses the conditional cross-correlation information to split one class into two sub-classes. The synopsis of this classification scheme is presented in figure 6. The criterion used to choose the class to be split into two sub-classes necessitates the calculation of the separability measure defined in (19). In each class the distance measure based on the cross-correlation information is calculated for each pixel from (14). The class X_m is temporarily split into two sub-classes X_{m1} and X_{m2} by comparing the value of the distance for each pixel with respect to the mean over the whole class equal to $\underline{d}_m = \frac{1}{N_m} \sum_{N_m} d_2(\mathbf{A}_{12} | \mathbf{A}_{22}, X_m)$ for each $\mathbf{A} \in X_m$.

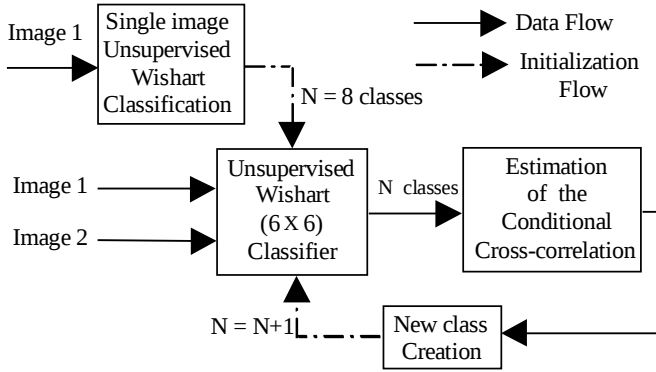


Figure 6. Synopsis of the dual polarimetric SAR data classification procedure using the polarimetric cross-correlation matrix probability.

The decision rule is defined by

$$\begin{aligned} &\text{For every } \mathbf{A} \in X_m, \text{ if } d_2(\mathbf{A}_{12} | \mathbf{A}_{22}, X_m) > \underline{d} \\ &\text{then } \mathbf{A} \in X_{m1}, \text{ else } \mathbf{A} \in X_{m2} \end{aligned} \quad (20)$$

The class to be split X_s is the one presenting the most distant sub-classes and verifies

$$Sp(X_{s1}, X_{s2}) > Sp(X_{m1}, X_{m2}) \quad \forall s \neq m \quad (21)$$

where $Sp(X_{s1}, X_{s2})$ represents the separability between classes X_{s1} and X_{s2} and is defined in (19).

The classification is initialized with the result of the single image based on the Wishart classification procedure. The number of classes is iteratively increased till a termination criterion is met

The classification algorithm is defined as follows

- Step 1: Initialize the 8 class distribution from the unsupervised Wishart classification on one of the separate polarimetric data sets. The number of classes, N , is equal to 8.
- Step 2: Apply the N class unsupervised dual data sets Wishart classification using the (6×6) polarimetric representation, till a termination criterion is met.
- Step 3: If a general termination criterion is met, go to Step 6.
- Step 4: For each class, perform a temporary splitting into two sub-classes, by applying the criterion defined in (20).

Step 5: Effectively split the class verifying (21), $N = N + 1$, go back to Step 2.

Step 6: Stop.

Similarly to the dual image classification scheme, the general termination criterion may be obtained by evaluating the classification global quality from the parameter defined in [19], or by fixing a maximum number of classes. This procedure is less computationally intensive than the former one since the number of classes remains inferior or equal to the final one.

4.2.2 Application to POLSAR Data

The segmentation of P - L and P - C band dual data sets in 16 clusters leads to almost similar results using both dual classification methods. The distribution of the classes is slightly different, but leads to an equivalent interpretation of the forest parcels. Due to the merging procedure, the classes produced by the first method are slightly more compact.

We apply both classification methods to the dual P - C band data set with a number of classes equal to 12.

The maximum likelihood classification scheme provides a segmentation of the observed scene which gathers the almost totality of the tree covered parcels two classes, classes number 2 and 4, as can be seen in figure 7.

During the class number reduction process, the polarimetric classes corresponding to the various types of forest show a low separability, compared to the point targets discriminated by C band data, and are merged in a single class so that the remaining clusters describe the bare soil areas. The joint use of these frequency bands with this classification method, using a small number of classes, does not provide good results for forest classification.

The classification based on the conditional probability of the cross-correlation matrix is initialized with 8 clusters resulting from the Wishart iterative classification applied on one of the separate data sets and the class splitting procedure is run till the number of classes reaches 12.

The classification results are shown in figure 8.

The classification provides better results than these depicted in figure 7 and provide good global information about the observed scene. The different types of forest parcels can be distinguished and the bare

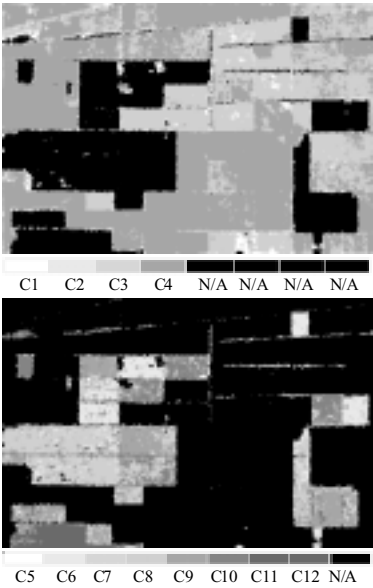


Figure 7. Results of the unsupervised P and C band dual data classification with 12 classes.

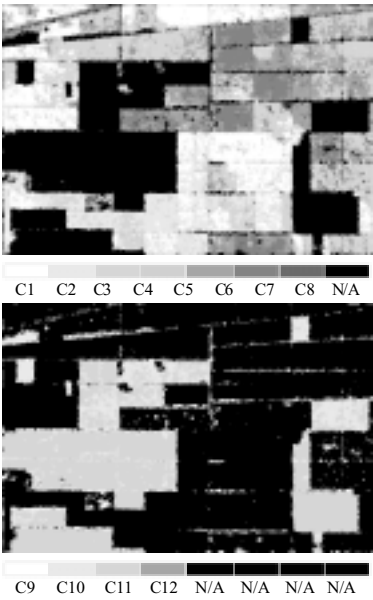


Figure 8. Results of the unsupervised P and C band dual data classification using the polarimetric cross-correlation matrix probability, with 12 classes.

soil areas as well as the major part of the ways are discriminated. This method is an efficient alternative to the maximum likelihood dual classification, when the reduction procedure may merge close classes. An important gain in computation time was observed too.

5. DUAL POLARIMETRIC CLASS DESCRIPTION

Once dual data sets are segmented according to their full polarimetric characteristics, each resulting class may be analyzed by determining the average change in polarimetric properties from one image to the other. Variations of the mean backscattering phenomenon properties may be characterized by the following indicators :

- The total polarimetric power given by the $\text{span} = \lambda_1 + \lambda_2 + \lambda_3$ [10]
- The normalized eigenvalue spectrum defined by the entropy and the anisotropy [10].
- The average normalized backscattering mechanism described by $\underline{\alpha}$, $\underline{\beta}$, $\underline{\delta}$, and $\underline{\gamma}$ [10].

The variation of the total power of the eigenvalue spectrum may be described by differences or ratios, while the change of backscattering mechanism necessitates the use of special unitary operators.

5.1 Special Unitary Transformation

Any pure target coherency matrix may be decomposed as follows

$$\mathbf{T} = \lambda(\mathbf{V}\mathbf{\Sigma}_0\mathbf{V}^\dagger) = \lambda\mathbf{u}\mathbf{u}^\dagger \text{ with } \mathbf{\Sigma}_0 = \text{diag}[1 \ 0 \ 0] \quad (22)$$

where $\mathbf{V}$ and $\mathbf{\Sigma}_0$ respectively stand for the eigenvector and normalized eigenvalue matrix of $\mathbf{T}$. $\mathbf{u}$ is the unitary eigenvector related to the single non-zero eigenvalue λ and represents the normalized target vector. The constant structure of the normalized eigenvalue matrix involves that coherency matrices measured at different frequencies $\mathbf{T}_1$ and $\mathbf{T}_2$ present normalised target vectors, $\mathbf{u}_1$ and $\mathbf{u}_2$, which are linked by the way of a special unitary transformation as shown in (23).

$$\mathbf{V}_2\mathbf{\Sigma}_0\mathbf{V}_2^\dagger = \mathbf{U}_3(\mathbf{V}_1\mathbf{\Sigma}_0\mathbf{V}_1^\dagger)\mathbf{U}_3^\dagger \Rightarrow \mathbf{u}_2 = \mathbf{U}_3\mathbf{u}_1 \quad (23)$$

where $\mathbf{U}_3$ is a (3×3) complex special unitary operator verifying $\mathbf{U}_3^{-1} = \mathbf{U}_3^\dagger$ and $|\mathbf{U}_3| = +1$. This operator completely defines the change of scattering basis from $\mathbf{T}_1$ to $\mathbf{T}_2$ and then summarizes the

$$\begin{aligned}
\mathbf{G}_1 &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \mathbf{G}_2 &= \begin{bmatrix} 0 & -j & 0 \\ j & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \mathbf{G}_3 &= \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \mathbf{G}_4 &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\
\mathbf{G}_5 &= \begin{bmatrix} 0 & 0 & -j \\ 0 & 0 & 0 \\ j & 0 & 0 \end{bmatrix} & \mathbf{G}_6 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} & \mathbf{G}_7 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -j \\ 0 & j & 0 \end{bmatrix} & \mathbf{G}_8 &= \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1/\sqrt{3} & 0 \\ 0 & 0 & -2/\sqrt{3} \end{bmatrix} \\
\mathbf{A}_1 &= \begin{bmatrix} \cos(w_1) & j \sin(w_1) & 0 \\ j \sin(w_1) & \cos(w_1) & 0 \\ 0 & 0 & 1 \end{bmatrix} & \mathbf{A}_2 &= \begin{bmatrix} \cos(w_2) & \sin(w_2) & 0 \\ -\sin(w_2) & \cos(w_2) & 0 \\ 0 & 0 & 1 \end{bmatrix} & \mathbf{A}_3 &= \begin{bmatrix} e^{jw_3} & 0 & 0 \\ 0 & e^{-jw_3} & 0 \\ 0 & 0 & 1 \end{bmatrix} & \mathbf{A}_4 &= \begin{bmatrix} \cos(w_4) & 0 & j \sin(w_4) \\ 0 & 1 & 0 \\ j \sin(w_4) & 0 & \cos(w_4) \end{bmatrix} \\
\mathbf{A}_5 &= \begin{bmatrix} \cos(w_5) & 0 & \sin(w_5) \\ 0 & 1 & 0 \\ -\sin(w_5) & 0 & \cos(w_5) \end{bmatrix} & \mathbf{A}_6 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(w_6) & j \sin(w_6) \\ 0 & j \sin(w_6) & \cos(w_6) \end{bmatrix} & \mathbf{A}_7 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(w_7) & \sin(w_7) \\ 0 & -\sin(w_7) & \cos(w_7) \end{bmatrix} & \mathbf{A}_8 &= \begin{bmatrix} e^{jw_8/\sqrt{3}} & 0 & 0 \\ 0 & e^{jw_8/\sqrt{3}} & 0 \\ 0 & 0 & e^{j-2w_8/\sqrt{3}} \end{bmatrix}
\end{aligned}$$

Figure 9. The eight Gell-Mann SU3 generators $\mathbf{G}_i$ and their corresponding operators $\mathbf{A}_i$.

modification in the scattering mechanism as the observation frequency varies.

As an element of the special unitary matrix group SU3, $\mathbf{U}_3$ may be expressed in terms of a matrix complex exponential of a linear combination of eight generators which form a basis matrix set [25, 26].

$$\mathbf{U}_3 = \exp \left(j \sum_{i=1}^8 w_i \mathbf{G}_i \right) \quad (24)$$

where w_i represent a real scalar angular variable and $\mathbf{G}_i$ the corresponding generator. The Gell-Mann matrix set is commonly used as a basis for SU3. It had applications in particles physics theory [26] and was introduced in the important case of backscatter of polarized electromagnetic waves in [27].

In Fig. 9 are given the eight traceless hermitian Gell-Mann generators $\mathbf{G}_i$ and their corresponding SU3 operators $\mathbf{A}_i = \exp(jw_i \mathbf{G}_i) = \mathbf{U}_3(w_i)$.

The operators $\mathbf{A}_2$, $\mathbf{A}_5$ and $\mathbf{A}_7$ correspond to rotations around the different components of the unitary target vector, while $\mathbf{A}_1$, $\mathbf{A}_4$ and $\mathbf{A}_6$ represent the corresponding elliptic transformations. $\mathbf{A}_3$ and $\mathbf{A}_8$ modify the phase difference between the components of $\mathbf{u}$. Three of these operators, $\mathbf{A}_1$, $\mathbf{A}_4$ and $\mathbf{A}_7$, have equivalent representations in the (2×2) SU2 group of the special unitary operators dedicated to the transformation of the scattering matrix $\mathbf{S}$. The scalar variables w_1 , w_4 , w_7 are proportional to respectively v the bouncing angle, τ the symmetry angle, and ϕ , the orientation angle as defined by Huynen in the phenomenological analysis of the target polarization fork [8]. The characterization of the polarimetric transformation of a normalized target vector consists in determining the complete set of 8

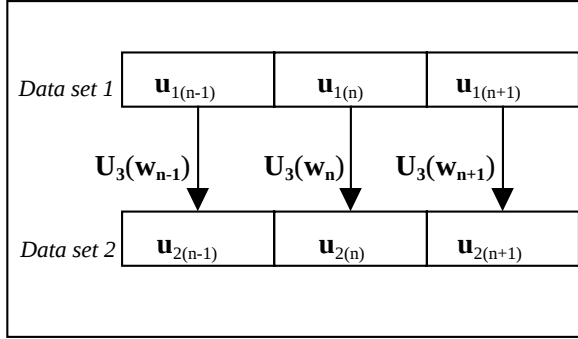


Figure 10. Special unitary transformations between successive dual frequency samples.

real Gell-Mann parameters, $\mathbf{w} = [w_1, \dots, w_8]$, which exactly defines a special unitary operator.

The resolution of the non-linear relation given by $\mathbf{u}_2 = \mathbf{U}_3(\mathbf{w})\mathbf{u}_1$, leads to an under-determined system of five observables with eight variables and has an infinite number of solutions [28]. A performing method to extract the Gell-Mann parameter vector $\mathbf{w}$ consists in assuming that the polarimetric variation from one data set to the other remains constant $\mathbf{U}_3(\mathbf{w}_{n+1}) \approx \mathbf{U}_3(\mathbf{w}_n)$ over two sample periods as shown in Fig. 10.

This assumption permits to obtain an over-determined system of equations whose resolution is performed by the way of a least-square non-linear optimization technique aiming to determine $\mathbf{w}$ which minimizes the real scalar ε^2 as shown in (25).

$$\varepsilon^2(\mathbf{w}) = \mathbf{K}^\dagger \mathbf{K} \quad \text{with} \quad \mathbf{K} = \begin{bmatrix} \mathbf{u}_{2(n)} - \mathbf{U}_3(\mathbf{w})\mathbf{u}_{1(n)} \\ \mathbf{u}_{2(n+1)} - \mathbf{U}_3(\mathbf{w})\mathbf{u}_{1(n+1)} \end{bmatrix} \quad (25)$$

5.2 Application to POLSAR Data

The method used to minimize the least square error is based on the Levenberg-Marquardt algorithm which gathers the advantages of both the conjugate gradient and Newton algorithms. It requires the calculation of the derivative of the error with respect to each of the Gell-Mann parameters.

We applied the error minimization algorithm to the classified dual data set obtained with P and L band data. Each dual cluster is then

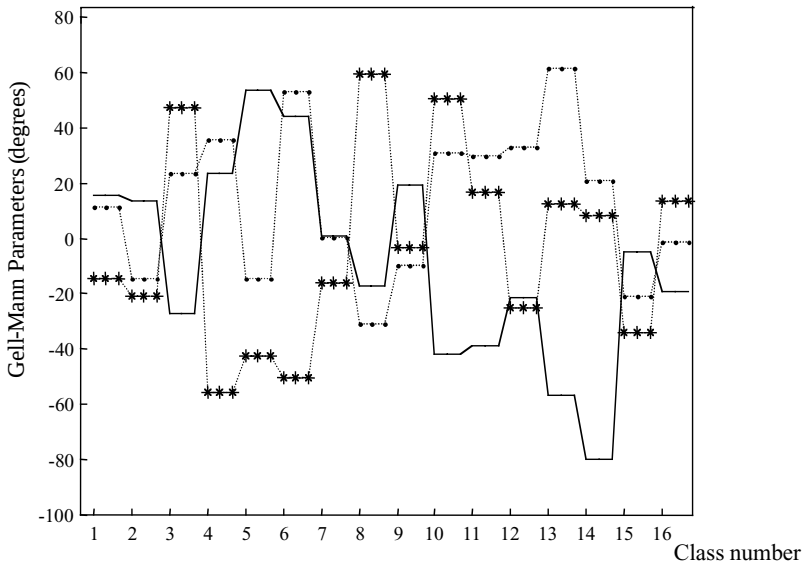


Figure 11. Gell-Mann parameters (degrees), w_2 (solid line), w_5 (dashed dot line), w_8 (dashed star line) over the 16 classes obtained with P and L band data sets.

represented by a set of eight average Gell-Mann parameters. In Fig. 11 are represented the values of three of these parameters for the sixteen classes.

The sixteen dual classes are characterized by vectors of mean Gell-Mann parameters which correspond to transformations of the mean scattering mechanism from one image to the other.

It is interesting to note that each cluster has a different mean Gell-Mann parameters set. This particularity will be used in future studies to develop an interpretation of the change in polarimetric properties occurring within each dual class.

6. CONCLUSION

In this paper, we introduced a new classification scheme for dual frequency polarimetric SAR data sets using a (6×6) polarimetric coherency matrix to simultaneously take into account the full polarimetric information from both images. Two classification methods were proposed. The first one was based on an iterative algorithm using a maximum likelihood decision rule evaluated from the Wishart density

function of the (6×6) matrix. The initialization of this classification is realized with a combination of the H - α classification results from each image providing 64 initial classes. Once the iterative algorithm has converged, a class number reduction technique is applied to improve the representation of each class characteristics. The results obtained with this classification show an important improvement in the description of the different types of natural media encountered in a forest scene. Parcels containing different types of trees can be distinguished and small classes such as roads and small forest parcels are discriminated. The class number reduction technique enhances the class compactness and improves the interpretation possibilities.

This reduction procedure may, in case of point targets, lead to the merging of large areas into small number of polarimetric classes. In order to overcome this problem, a second technique is proposed which introduces the polarimetric cross-correlation information and refines the results by iteratively creating new classes during the classification.

Once dual data sets are classified, the analysis of the resulting clusters is realized by determining the rigorous change in polarimetric properties from one image to the other. The polarimetric variations are parameterized by eight real coefficients derived from the decomposition of a special unitary operator on the Gell-Mann basis. Each dual class is thus characterized by a set of eight average real Gell-Mann coefficients.

REFERENCES

1. Rignot, E., R. Chellappa, and P. Dubois, "Unsupervised segmentation of polarimetric SAR data using the covariance matrix," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30, No. 4, 697–705, July 1992.
2. Zebker, H. A., J. J. van Zyl, S. L. Durden, and L. Norikane, "Calibrated imaging radar Polarimetry: techniques examples and applications," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 29, 942–961, 1991.
3. Hara, R. G. Atkins, S. H. Yueh, R. T. Shin, and J. A. Kong, "Application of neural networks to radar image classification," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, 100–110, January 1994.
4. van Zyl, J. J. and C. F. Burnette, "Bayesian classification of polarimetric SAR images using adaptive a-priori probabilities," *International Journal of Remote Sensing*, Vol. 13, 835–840, 1992.

5. Pottier, E., "On full polarimetric target decomposition theorems with application to classification and identification of real target cross section," proceedings of *International Radar Conference*, 330–335, Paris, May 1994.
6. van Zyl, J. J., "Unsupervised classification of scattering behavior using radar polarimetry data," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 27, 36–45, 1989.
7. Boerner, W. M. et al., "Polarimetry in radar remote sensing: basic and applied concepts," Chapter 5 *Principles and Applications of Imaging Radar, The Manual of Remote Sensing*, 3rd Edition, The American Society for Photogrammetry and Remote Sensing, March 1998.
8. Huynen, J. R., "Phenomenological theory of radar targets," Ph.D. dissertation, Drukkerij Bronder-offset, N. V. Rotterdam, 1970.
9. 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, September 1995.
10. 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. 35, No. 1, 68–78, January 1997.
11. Freeman, A. and S. Durden, "A three component scattering model to describe polarimetric SAR data," *SPIE*, Vol. 1748, 213–225, Radar Polarimetry, 1992.
12. Dong, Y., B. Forster, and C. Ticehurst, "A new decomposition of radar polarization signatures," *IEEE Trans. Geoscience and Remote Sensing*, Vol. 36, No. 3, 933–939, 1998.
13. Shi, J. and J. Dozier, "On estimation of snow water equivalence using SIR-C/X-SAR," proceedings of the *Second International Workshop on Retrieval of Bio- and Geo-physical Parameters from SAR Data for Land Application*, Noordwijk, The Netherlands, October 1998.
14. Floricioiu, D. M., "Polarimetric signatures and classification of alpine terrain by means of SIR-C/X-SAR," Ph.D. dissertation, Innsbruck, Austria, 1997.
15. Chen, K. S. et al., "Classification of multifrequency polarimetric SAR, image using a dynamic learning neural network," *IEEE Trans. Geoscience and Remote Sensing*, Vol. 34, No. 3, 814–820, 1996.
16. Freeman, A., S. Durden, and R. Zimmerman, "Mapping Sub-Tropical vegetation using Multi-Frequency Multi-Polarization

- SAR data," Proceedings of *IGARSS*, 1686–1689, Houston, USA, June 1992.
17. 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.
 18. Kong, J. A., S. H. Yueh, R. T. Shin, and J. J. van Zyl, "Classification of earth terrain using polarimetric synthetic aperture radar images," Chapter 6 in *PIER*, Vol. 3, J. A. Kong (Ed.), Elsevier 1990.
 19. Lee, J. S., M. R. Grunes, T. L. Ainsworth, L. Du, D. L. Schuler, and S. R. Cloude, "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, Sept. 1999, also proceedings of the *Fourth International Workshop on Radar Polarimetry*, PIERS 1998, 13–17, Nantes, France, July 1998.
 20. Cloude, S. R. and K. P. Papathanassiou, "Polarimetric SAR interferometry," *IEEE Trans. Geoscience and Remote Sensing*, Vol. 36, No. 5, 1551–1565, 1998.
 21. Ferro-Famil, L., E. Pottier, J. P. Dedieu, C. Corgier, and J. Sallard, "Application of polarimetric SAR data processing to snow cover remote sensing. Validation using optical images and ground data," proceedings of the *Committee on Earth Observing Satellites SAR Workshop*, 26–29, CNES, Toulouse, France, October 1999.
 22. Goodman, N. R., "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *Ann. Math. Statist.*, Vol. 34, 152–177, 1963.
 23. Muirhead, R. J., *Aspects of Multivariate Statistical Theory*, John Wiley and Sons, New-York, ISBN 0-471-094442-0.
 24. Pottier, E. and J. S. Lee, "Application of the H/A/ α polarimetric decomposition theorem for unsupervised classification of fully polarimetric SAR data based on the wishart distribution," proceedings of the *Committee on Earth Observing Satellites SAR Workshop*, 26–29, CNES, Toulouse, France, October 1999.
 25. Cloude, S. R., "Group theory and polarization algebra," *OPTIK*, Vol. 75, No. 1, 26–36, January 1986.
 26. Joshi, A. W., *Elements of Group Theory for Physicists*, Wiley Eastern Limited, New Delhi, September 1988.
 27. Cloude, S. R., "Lie groups in electromagnetic wave propagation and scattering," *Journal of Electromagnetic Waves and Applications*, Vol. 6, No. 8, 947–974, 1992.

28. Cloude, S. R. and E. Pottier, "Matrix difference operators as classifiers in polarimetric radar imaging," *L'Onde Electrique*, Vol. 74, No. 3, 34-40, May-June 1994,