Bias Compensation in $H/A/\alpha$ Polarimetric SAR Decomposition and Its Implication for the Classification

Mohamed Yahia*, **Faten Khalfa, Marwa Chabir, and Taoufik Aguili**

Abstract—Classification of land cover types is one important application of polarimetric synthetic aperture radar (PolSAR) remote sensing. There are numerous features that can be extracted from PolSAR images. Among them, eigenvalues λ_i, entropy H, alpha angle α, and anisotropy A are effective and popular tools for the analysis and quantitative estimation of the physical parameters. Nevertheless, the speckle noise appearing in PolSAR images reduces the accuracy of image classification. Consequently, it should be filtered correctly. Generally, filtering PolSAR data generate biased estimates of $\lambda_i/H/A/\alpha$ parameters. In this paper, we studied the effects of bias compensation on supervised and unsupervised PolSAR image classification. We applied the asymptotic quasi maximum likelihood estimator AQ-MLE and Yahia/Aguili's bias compensation methods. To improve the classification accuracies, we demonstrated that bias compensation must be associated with speckle reduction. The combination of the *span* with biased parameters reduced the effects of bias but did not eliminate it totally. Simulated and real data were used for validation.

1. INTRODUCTION

Synthetic Aperture Radar (SAR) is a powerful remote sensing tool for the observation and characterization of the Earth's surface. Originally, SAR systems produced single complex SAR images. However, the actual SAR systems produce different SAR images of the same scene under some type of diversity. In SAR Interferometry (InSAR) [1], different SAR images are acquired from slightly different positions, permitting the gain of information about the Earth's surface topography. In Polarimetric SAR (PolSAR) [1], images are acquired using the polarization diversity. Therefore, PolSAR data give information about the scattering properties of the imaged scene and the geometry of the scatterers on it. The importance of PolSAR data has augmented in the last years, especially after the launch of many spaceborne missions such as: Radarsat-2 (C-band), ALOS (L-band) and TerraSAR-X (X-band). For every resolution cell (i.e., pixel), PolSAR data are generally represented by the scattering or the covariance matrices.

Target Decompositions (TD) were developed to cluster the scattering or the covariance matrices into basic scattering mechanisms. The TD theorems can be divided into two categories: coherent target decompositions and incoherent target decompositions [2]. Among different approaches proposed in the literature, the incoherent eigendecomposition TD represents one appropriate tool to perform data interpretation of the distributed targets such as natural scatterers [2, 3]. Considering this approach, Cloude and Pottier [3] defined entropy H, anisotropy A and alpha α angle. Entropy and anisotropy are used to characterize the heterogeneity of the media scattering. Alpha angle is an indicator of the type of the scattering mechanism.

Classification of land cover types is one important application of PolSAR remote sensing. Various supervised and unsupervised techniques have been proposed. Many of the proposed techniques are

* Corresponding author: Mohamed Yahia (mohamed_yahia1@yahoo.fr).

The authors are with the Labo. SYSCOM, ENIT, BP. 37, Tunis le Belvédère 1002, Tunisia.

based on physical scattering mechanisms obtained from TD methods. The most popular one is the unsupervised classification technique proposed by Cloude and Pottier where H/α plane is divided into eight classes using arbitrarily fixed thresholds of the boundaries [3]. The divided H/α plane is hybridized with other classifier such as fuzzy c-means [4], support vector machines [5, 6], neural networks [7] and Wishart classifier [8]. The last method was extended by introducing the anisotropy parameter A leading to 16 initial classes [9]. The $H/A/\alpha$ parameters are able to describe the scattering mechanisms within the resolution cell. Nevertheless, they do not contain information about the backscattering power. To overcome this limitation, Yahia and Belhadj combined the diagonal elements the coherency matrix and the $H/A/\alpha$ parameters to train the multilayer perceptron (MLP) neural network (NN) classifier [10]. The backscattering power (*span*) and $H/A/\alpha$ parameters were combined to initialize an unsupervised complex Wishart classifier [11, 12]. In [13], Zhang et al. reduced the dimension of the $span/H/A/\alpha/$ feature vector by using the principal component analysis to train a MLP NN. Since the eigenvalues of Cloude-Pottier decomposition express the intensities of scattering mechanisms, they contain both scattering mechanism and intensity information. Gou et al. exploited the eigenvalues probability density functions (pdf) for Bayesian classification [14]. Thus, eigenvalues and $H/A/\alpha$ parameters are widely used as feature vectors for various supervised and unsupervised classifiers.

Nevertheless, due to the coherent record of radar signals, SAR data are affected by the speckle noise which is one of the major problems of the SAR imagery [15]. Therefore, for best image interpretation, speckle noise must be reduced [16, 17]. This can be achieved by averaging the covariance matrices or coherency matrices from neighboring pixels. This operation is known as multilook processing. However, overaveraging degrades the spatial resolution, and insufficient averaging generates biased estimates [18]. Diverse bias removal procedures have been proposed in the literature. The asymptotic quasi maximum likelihood estimator (AQ-MLE) of sample eigenvalues is derived from the eigenvalue pdfs [18]. Foucher et al. [19] introduced a bootstrap technique to estimate eigenvalues, entropy, anisotropy, and α angle. The bootstrap on eigenvalues gives similar bias correction to AQ-MLE [19]. Lee et al. [20] exploited the linear dependence of the entropy bias on the entropy to its bias removal. This technique requires the value of entropy for an infinite number of looks. This shortcoming was surmounted in [21]. However, the main weaknesses of the previous techniques are their dependency on a precise estimation of the independent averaging samples and their amplification of the noise variances [22]. Recently, Yahia and Aguili theoretically demonstrated that the means and variances of the eigendecomposition can be approximated by a linear rule [22]. Then, a linear regression was applied to compensate the biases. In addition to its versatility, independency of the knowledge of the equivalent number of looks (ENL) and high ability for bias compensation, Yahia/Aguili's approach reduced the variance of noises on the eigendecomposition parameters [22] which is favorable for target classification. In fact, Lee et al. studied the implication of speckle filtering for the classification [23]. They found that speckle filtering of the intensity images increased the performances of the classification. However, it is important to notice that speckle filtering acts on the intensity images whereas bias compensation is a succeeding stage that acts on the eigendecomposition parameters.

To our knowledge, the impact of bias compensation on the classification has not been investigated quantitatively yet. López-Martínez et al. stated [18]: "*Nevertheless, the biases are no so critical for classification applications, since in these cases, the effects of the biases can be mitigated by the fact that other parameters, such as the span, are also considered*". However, many classification schemes used only biased parameters as feature vectors such as H/α [2–9, 13], eigenvalues [14], etc. In these cases, the effects of the biases cannot be mitigated with other parameters. In addition, when biased parameters are combined with other parameters such as the *span*, it is important to quantify the mitigation of the effects of biases.

In this paper, we studied the implication of bias elimination for supervised and unsupervised classification. We selected H/α, $\lambda_1/\lambda_2/\lambda_3$ and $span/H/A/\alpha$ as feature vectors. Particularly, in the last case, we studied the mitigation of the biases in $H/A/\alpha$ by the use of the *span* parameter. This paper is organized as follows. Section 2 introduces SAR polarimetry and the eigendecomposition of the coherency matrix. In Section 3, we introduce the real and simulated PolSAR data used for validation. Section 4 characterizes the multilook effects on eigendecomposition parameters. Section 5 deals with bias elimination. In Section 6, we demonstrate the impacts of bias elimination on the classification. Finally, Section 7 gives the conclusions of this paper.

2. $H/A/\alpha$ POLARIMETRIC TARGET DECOMPOSITION

PolSAR data are generally represented by $2 \times 2 S$ scattering matrix

$$S = \left[\begin{array}{cc} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{array} \right] \tag{1}$$

The subindices h and v represent the horizontal and vertical orthogonal polarizations, respectively. In the general monostatic backscattering configuration, we have $S_{hv} = S_{vh}$. The scattering matrix can be expressed in a vector Pauli basis:

$$k = \frac{1}{\sqrt{2}} \left[S_{hh} + S_{vv}, \, S_{hh} - S_{vv}, \, 2S_{hv} \right]^T \tag{2}$$

where ^{T} indicates transposition. The coherency matrix is:

$$T = E \left\{ k \cdot k^{*t} \right\} \tag{3}$$

where $E\{\cdot\}$ indicates the statistical expectation, and * denotes complex conjugation. T is 3×3 Hermitian positive define matrix. Its eigendecomposition is [3]

$$T = U \cdot \Lambda \cdot U^{*T} \tag{4}$$

$$\Lambda = \left[\begin{array}{ccc} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{array} \right]_{\lambda_1 > \lambda_2 > \lambda_3} \tag{5}$$

λ_i for $i = 1$, 2, 3, represent the true eigenvalues of T [3]. The 3×3 complex unitary matrix U contains the eigenvectors.

$$u_i = \left[\cos(\alpha_i), \, \sin(\alpha_i) \cos(\beta_i) e^{j\delta_i}, \, \sin(\alpha_i) \sin(\beta_i) e^{j\gamma_i} \right]. \tag{6}$$

The $H/A/\alpha$ parameters are defined as follows [3]:

$$H = \sum_{i=1}^{3} p_i \log_3 (p_i) \quad p_i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \lambda_3} \tag{7}$$

$$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3} \tag{8}$$

$$\alpha = \sum_{i=1}^{3} p_i \alpha_i. \tag{9}$$

The coherency matrix T is usually estimated by means of a spatial averaging as

$$\hat{T} = \frac{1}{N} \sum_i = 1^N k_i k_i^{*T} \tag{10}$$

where $\hat{T}$ is the sample coherency matrix and N the number of independent samples included in the average which is also known as the number of looks. The eigendecomposition of T is estimated through the eigendecomposition of $\hat{T}$ which defines the sample parameters $\hat{\lambda}_i$, $\hat{H}$, $\hat{A}$ and $\hat{\alpha}$.

3. SIMULATED AND REAL DATA

Both simulated and real data were used in this paper. The real test site was a PolSAR data of Les-Landes, France (Fig. 1). It was a single-look complex scattering matrix acquired by NASA/JPL polarimatric AIRSAR and was characterized by rectangular forested areas having different ages [9]. A, B, C, D and E zones were used for validation (see Fig. 1). We performed a Monte Carlo study using an extensive data set containing 40 simulated coherency matrices obtained by the simulation procedure used in [22]. For validation, we considered eight simulated targets (i.e., coherency matrices) $T_1, T_2, \ldots, T_8$ (see Table 1). We selected one target from each scattering mechanism in Fig. 2(b) [3].
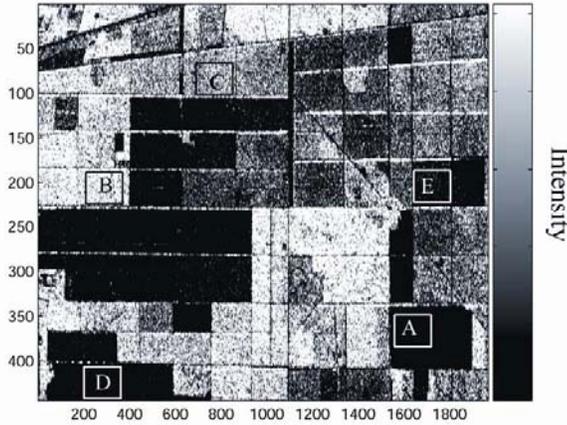
Figure 1. *span* image $(S_{hh}^2 + 2 * S_{hv}^2 + S_{vv}^2)$ of the test site "Les Landes". Zones in boxes are the real testing areas.

Table 1. Simulated coherency matrices used for validation.

	T_{11}	T_{12}	T_{13}	T_{22}	T_{23}	T_{33}
$T_1 \times 10^3$	0.23	$0.44 + 0.14i$	$-0.06 - 0.003i$	1.21	$-0.16 + 0.03i$	0.04
$T_2 \times 10^3$	0.97	$0.52 - 0.51i$	$0.28 - 0.34i$	0.69	$0.39 - 0.002i$	0.24
$T_3 \times 10^3$	0.38	$-0.006 + 0.006i$	$0.03 - 0.02i$	0.36	$-0.01 - 0.01i$	0.06
$T_4 \times 10^3$	0.009	$-0.04 + 0.03i$	$-0.004 + 0.02i$	0.65	$0.05 + 0.33i$	0.89
$T_5 \times 10^3$	0.34	$0.008 + 0.05i$	$-0.01 + 0.007i$	0.60	$-0.02 + 0.08i$	0.59
$T_6 \times 10^3$	0.18	$0.008 + 0.01i$	$-0.001 + 0.003i$	0.13	$0.0001 + 0.0001i$	0.18
$T_7 \times 10^3$	0.41	$-0.006 - 0.002i$	$-0.004 - 0.006i$	0.003	$-0.003 - 0.016i$	0.001
$T_8 \times 10^3$	0.41	$-0.08 - 0.06i$	$-0.036 - 0.02i$	0.11	$0.01 + 0.001i$	0.081

4. CHARACTERIZATION OF MULTILOOK EFFECTS ON THE EIGENDECOMPOSITION PARAMETERS

Multilook simulations for each target was processed for 3, 2×2, 3×3, 4×4, ..., 10×10 independent sample averages. Mean for each number of looks was computed using 2000 samples. We processed the testing areas $(A, \ldots, E)$ by 2×4, 3×6, 4×8 and 5×10 boxcar filter. As demonstrated in [18–22], we observe from Figs. 2(a) and (b) that sample entropy was always underestimated. For low entropy, bias was insignificant. Bias increased as the entropy increased. Sample anisotropy $\hat{A}$ was generally overestimated. Bias increased with the respect decrease of the anisotropy. For high anisotropy, bias was insignificant. For low anisotropy, bias was still important even after 10×10 looking. Regarding sample alpha angle $\hat{\alpha}$, for low values (e.g., surface scattering), the mean alpha angle was overestimated, for medium values (e.g., volume scattering), bias on $\hat{\alpha}$ was negligible, and for high values (e.g., urban areas), $\hat{\alpha}$ was underestimated. In conclusion, we remark that the biases decreased the dynamic ranges of the $H/A/\alpha$ parameters which was unfavorable for classification. Consequently, to increase the ability to discern different classes, bias elimination was essential.

Figure 3 plots the variances versus the means of the eigenvalues for different numbers of looks. As demonstrated in [18], we found that sample eigenvalue $\hat{\lambda}_1$ was always underestimated. Bias increased with the increase of the entropy. For low entropy, bias was not significant. Therefore, sample eigenvalue $\hat{\lambda}_3$ was always overestimated. Bias increased with the decrease of the anisotropy. The behavior of $\hat{\lambda}_2$ was
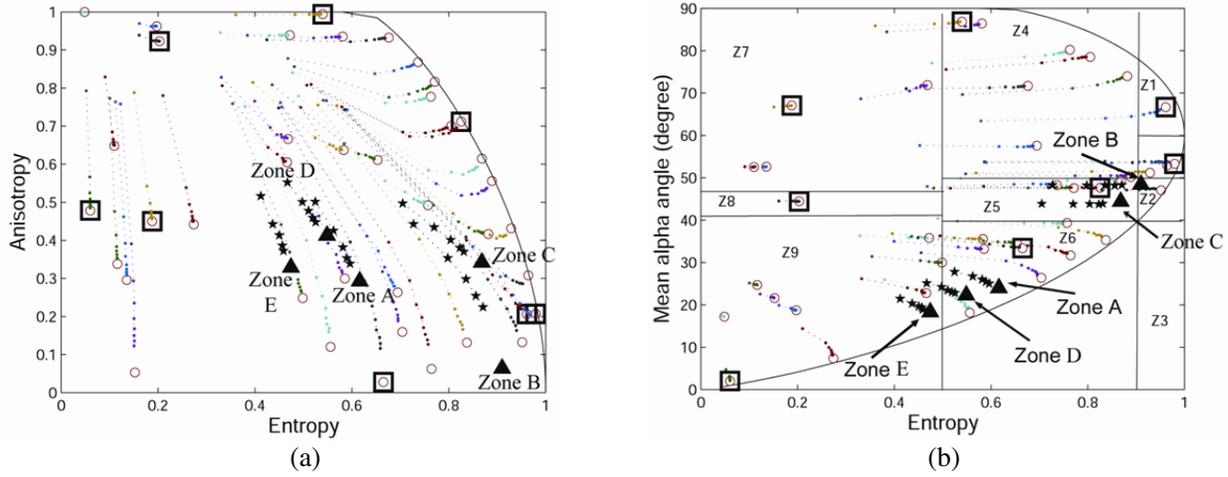
Figure 2. (a) H/A plane. Pointed and dashed lines: sample anisotropy versus sample entropy. (b) H/α plane. Pointed and dashed lines: sample alpha versus sample entropy. Simulated data (Circles: true values. Squares: true values of the simulated testing targets $T_1, T_2, \ldots, T_8$). Real data (Stars: multilooked values. Triangles: true values obtained by averaging the coherency matrices in the testing areas).
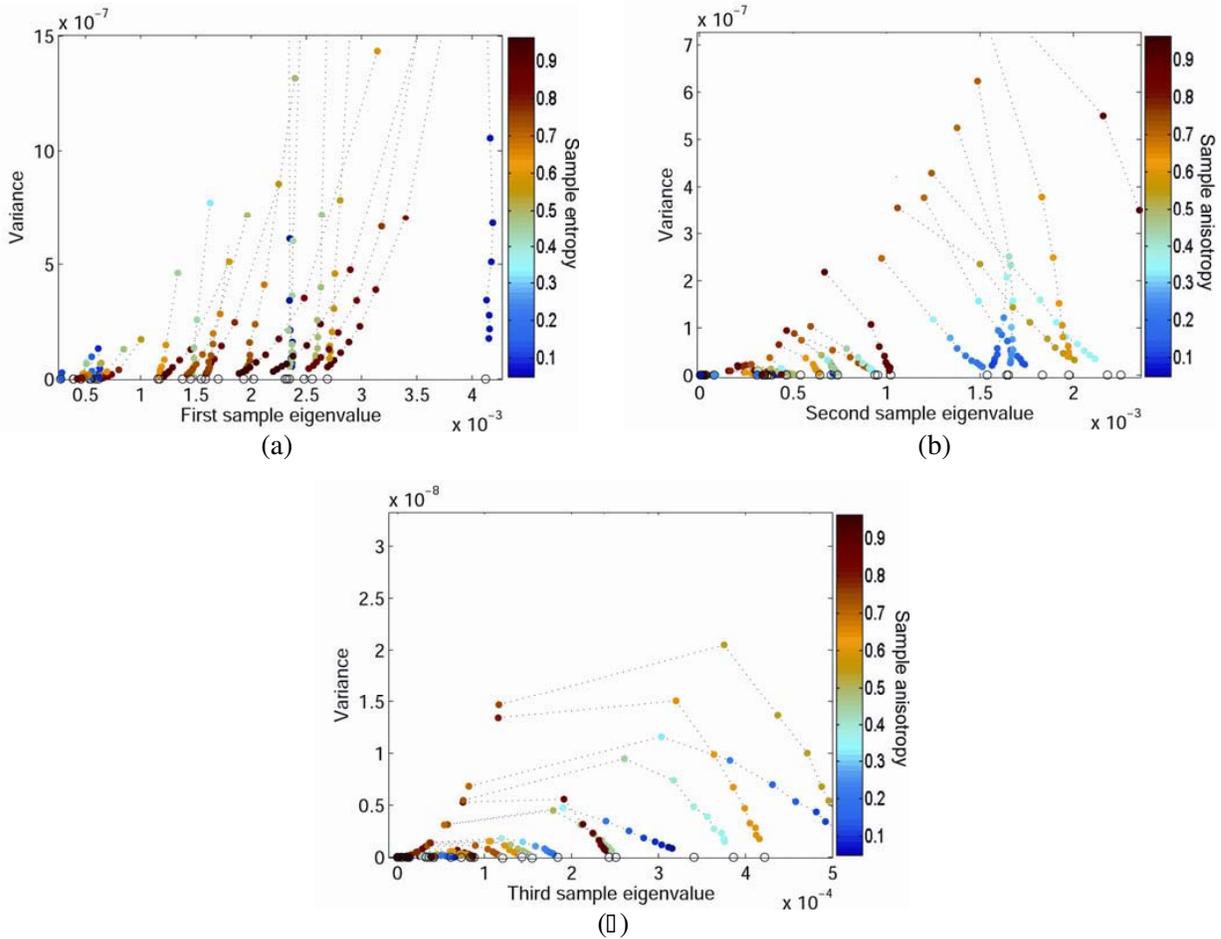
Figure 3. Variances versus means. Simulated data (Points: sample values. Circles: true values). Real data (Stars: multilooked values. Triangles: true values).

not predictable using a fixed number of looks. In fact, $\hat{\lambda}_2$ was generally underestimated, but for some cases it could be overestimated. In addition, it has been demonstrated that these biases depend not only on the number of averaged samples, but also on the difference between pairs of true eigenvalues [18, 21]. Since the behaviors of biases on eigenvalues are not simple as for $H/A/\alpha$ parameters, it seems that the bias elimination does not have an apparent effect on classification.

5. BIAS COMPENSATION PROCEDURES

The aim of bias compensation procedures is to shift the pdf of the biased parameter to obtain a mean equal to the true value. A variety of bias compensation procedures have been proposed in the literature [18–22]. However, these techniques reduced the bias but altered the variances of sample values. As established in [22], Yahia/Aguili's technique decreased the noise whereas the techniques proposed in [18–21] amplified it. Consequently, both bias and speckle variance influenced the performances of the classification. In this paper, we applied particularly the AQ-MLE [18] and Yahia/Aguili's method [22]. The AQ-MLE of sample eigenvalues are given by

$$\lambda_i = \hat{\lambda}_i - \frac{1}{N} \sum_{k=1,\, k\neq i}^{3} \frac{\hat{\lambda}_i \hat{\lambda}_k}{\hat{\lambda}_i - \hat{\lambda}_k} + O\left(N^{-1}\right) \tag{11}$$

In [22], Yahia and Aguili theoretically demonstrated that the means and variances of the eigendecomposition parameters for various numbers of looks can be approximated as:

$$\hat{X}_i = a_i \left(\operatorname{var}\left(\hat{X}_i\right)\right) + b_i \tag{12}$$

where X_i is an eigendecompostion parameter. a_i and b_i are two constants that can be estimated using first order interpolation for various numbers of looks. $\operatorname{var}(\cdot)$ is the variance.

In fact, López-Martínez et al. derived the expressions of means and variances of the sample eigenvalues as [21]:

$$\hat{\lambda}_i = \lambda_i + \frac{1}{N} \sum_{k=1,\, k\neq i}^{3} \frac{\lambda_i \lambda_k}{\lambda_i - \lambda_k} + O\left(N^{-2}\right) \tag{13}$$

$$\operatorname{var}\left(\hat{\lambda}_i\right) = \frac{\lambda_i^2}{N} + O\left(N^{-2}\right) \tag{14}$$

From Eqs. (14) and (15), we can derive the theoretic relation gathering means and variances of sample eigenvalues as

$$\hat{\lambda}_i = \lambda_i + \frac{\operatorname{var}\left(\hat{\lambda}_i\right)}{\lambda_i^2} \sum_{k=1,\, k\neq i}^{3} \frac{\lambda_i \lambda_k}{\lambda_i - \lambda_k} = \lambda_i + \operatorname{var}\left(\hat{\lambda}_i\right) \sum_{k=1,\, k\neq i}^{3} \frac{\lambda_k}{\lambda_i\left(\lambda_i - \lambda_k\right)} \tag{15}$$

Then

$$\hat{\lambda}_i = A_i \operatorname{var}\left(\hat{\lambda}_i\right) + B_i \tag{16}$$

where

$$A_i = \sum_{k=1,\, k\neq i}^{3} \frac{\lambda_k}{\lambda_i\left(\lambda_i - \lambda_k\right)} \tag{17}$$

and

$$B_i = \lambda_i \tag{18}$$

Equation (16) which shows a perfect linearity between means and variances of eigenvalues confirms the essence of the proposed bias compensation method.

Then, the Yahia/Aguili's method employed the linear relation (12) between biased sample values for various numbers of looks to reduce the bias on $\hat{\lambda}_i$, $\hat{H}$, $\hat{A}$ and $\hat{\alpha}$ [22]. The unbiased estimate of the parameter $\hat{X}_i$ is b_i. Table 2 displays the performances of the AQ-MLE and Yahia/Aguili method.

Table 2. Performances of the AQ-MLE and Yahia/Aguili's methods.

AQ-MLE	Yahia/Aguili
Require the estimation of the ENL	Do not require the estimation of the ENL
Require only the sample values for one number of looks	Require the sample values and the variances for various number of looks
Applied to sample eigenvalues only	Applied to all eigendecomposition parameters
Amplify the speckle noise	Reduce the speckle noise

To apply Yahia/Aguili's method, we averaged the simulated data of targets $T_1, T_2, \ldots, T_8$ using 4, 5, ..., 25 looks. After that, first order interpolation (12) was applied using these data [22]. Real data were processed by 7×7 boxcar filter. The ENL was about 25. The linear interpolation (12) was applied using 8, 9, ..., 49 sample averaging [22].

Figures 4(a), (b) and (c) show the initial and enhanced pdfs of the entropy, anisotropy and alpha angle of the studied zones (i.e., A, B, C, D and E) using Yahia/Aguili's method, respectively. We observe that the enhanced pdfs were moved to the true values, and the noise was decreased which was advantageous for classification. We also observe that the anisotropy was noisier than the entropy and alpha angle, but it offered extra information to discern between (A, D) and (B, C) classes. Figs. 4(d), (e) and (f) show examples of initial and enhanced pdfs of the eigenvalues using the AQ-MLE and

(a)

(b)

(c)

(d)
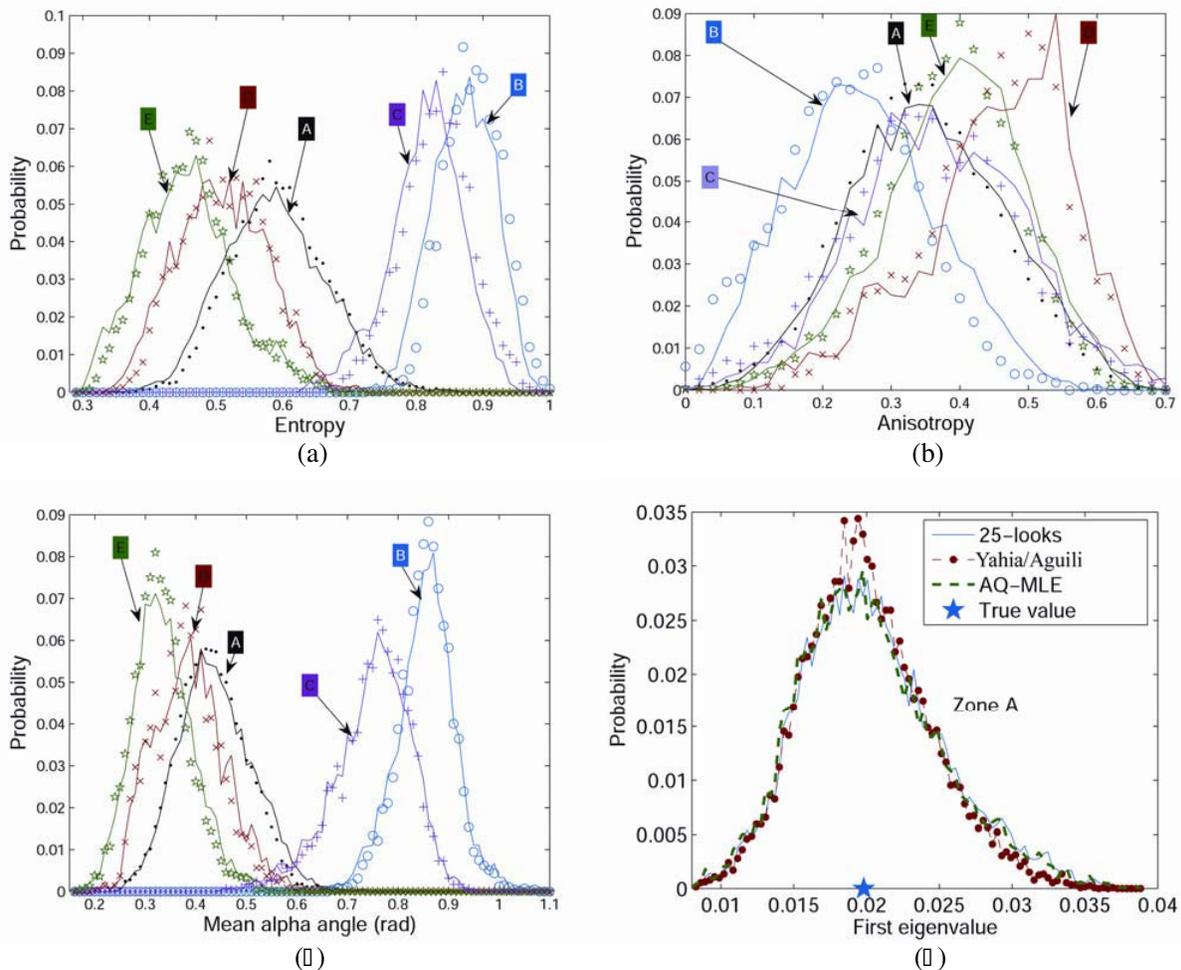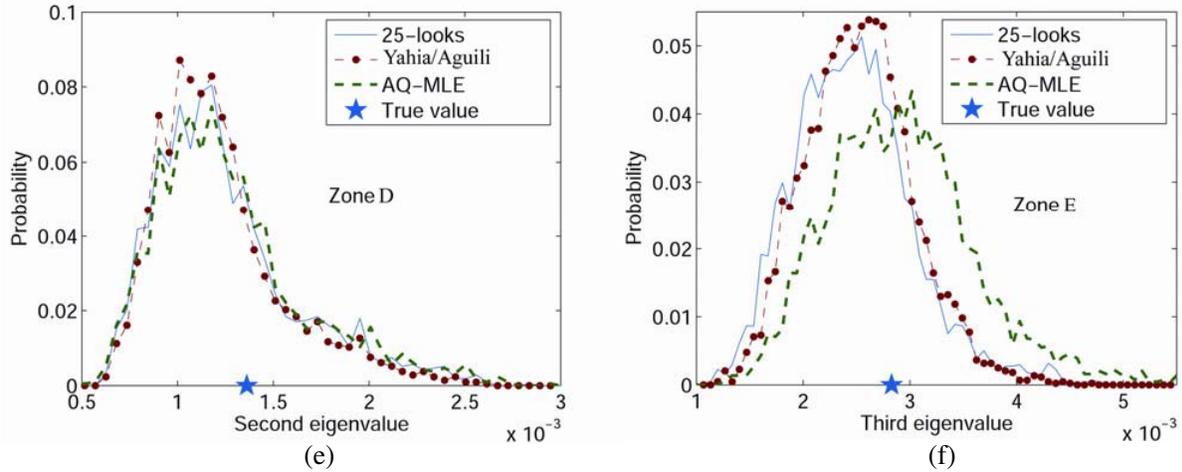
Figure 4. Original and compensated pdfs of sample eigendecomposition parameters.

Yahia/Aguili's method. The AQ-MLE extended the pdfs (i.e., amplified the speckle noise) whereas Yahia/Aguili's method constricted them (i.e., decreased the speckle noise).

6. IMPACTS OF BIAS COMPENSATION PROCEDURES ON THE CLASSIFICATION

In this section, we study the impacts of bias compensation procedures on the classification quantitatively. Various classification techniques have been proposed in the literature. Nevertheless, in this paper, we only concerned the effects of bias compensation on the classification and not the classifiers themselves. For this reason, we chose the basic forms of well-known supervised and unsupervised classification techniques that have been widely applied in the literature such as the H/α plane [3], fuzzy c-means (FCM) [24] and multilayer perceptron (MLP) [10, 13, 25]. The classifiers were initialized for both biased and unbiased inputs similarly. We used three feature vectors X_1, X_2 and X_3 as inputs where

$$X_1 = [\ H \quad \alpha\ ] \tag{19}$$

$$X_2 = [\ span \quad H \quad A \quad \alpha\ ] \tag{20}$$

$$X_3 = [\ \lambda_1 \quad \lambda_2 \quad \lambda_3\ ] \tag{21}$$

We selected feature vector X_1 to study the effects of bias on the classification using H/α plane. We used vector X_2 to study the mitigation of the biases in $H/A/\alpha$ by the use of the span parameter. Feature vector X_3 was employed to study the effects of bias compensation of eigenvalues (using the AQ-MLE and Yahia/Aguili's methods) on the classification.

6.1. Unsupervised Classification

6.1.1. H/α Plane

The H/α unsupervised classification scheme proposed by Cloude and Pottier has been widely applied to the analysis of PolSAR data. In fact, the principle of the classification is to divide the H/α plane into eight zones representing the main natural scattering mechanisms (see Fig. 2(a)). The main advantage of this unsupervised algorithm is that it provides terrain identification information where the most important kinds of scattering media can be discriminated. In the same figure, we plotted the means of sample $(\hat{H}, \hat{\alpha})$ points of the simulated data for various numbers of looks. We observe that the entropy bias has great effects on the classification especially for high values. Bias removal on entropy is crucial. Since the decision borders for α angle were situated in the region where this parameter was unbiased, it seems that bias on α angle does not change the classification results. However, from Fig. 5(a), we
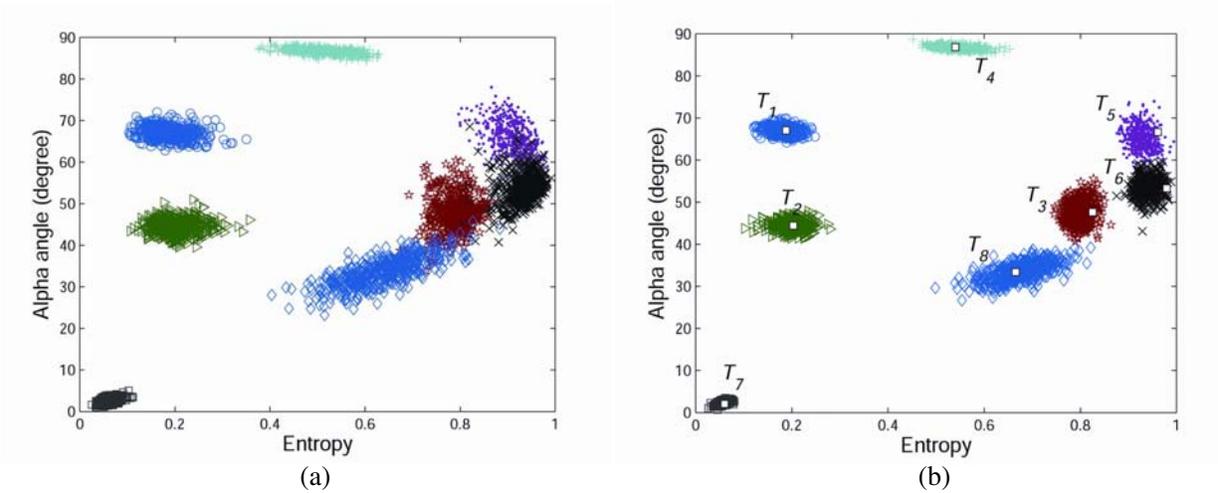
Figure 5. H/α plane using simulated data, (a) before bias compensation, (b) after bias compensation.
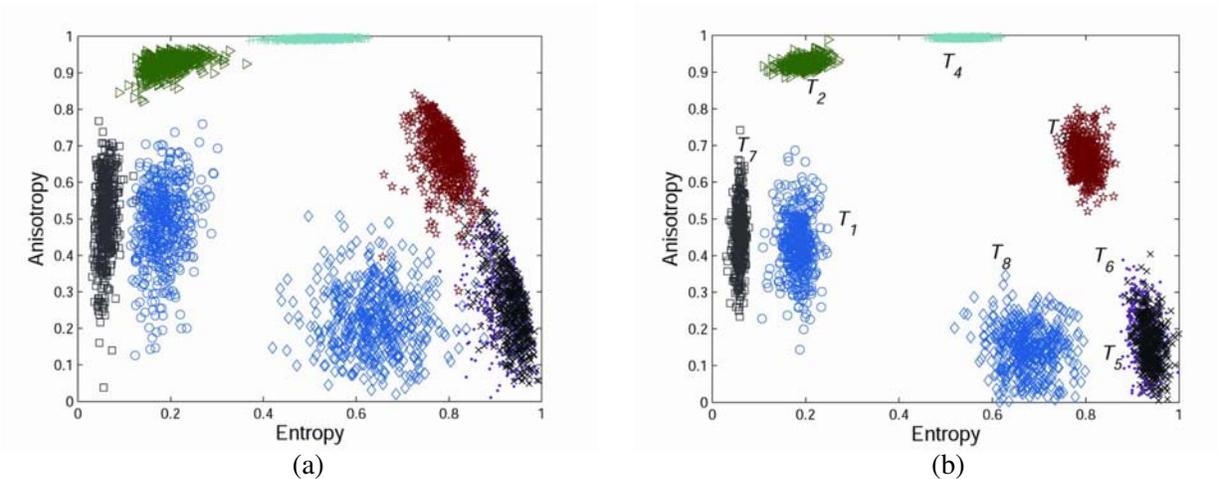
Figure 6. H/A plane using simulated data, (a) before bias compensation, (b) after bias compensation.

observe that the variance is important near the borders. Consequently, bias removal (with variance minimization) in α angle is essential too.

Figures 5(a) and 6(a) represent the original H/α and H/A planes, and Figs. 5(b) and 6(b) represent their compensated planes, respectively. We observe that Yahia/Aguili's method compensated the bias and reduced the noise. As a result, the overlapping between targets was reduced which was favorable for classification.

Table 3 displays the classification matrix of simulated data using Cloude and Pottier procedure. As observed, after bias compensation using Yahia/Aguili's method, all classification accuracies were increased significantly.

6.1.2. Fuzzy c-Means

The FCM algorithm, proposed by Bezdek et al. [24], is a partition based clustering algorithm. The standard FCM objective function of partitioning an image into c clusters is

$$J_m = \sum_{k=1}^{c} \sum_{i=1}^{n} (u_{ki})^m d^2 (x_i, \nu_k) \tag{22}$$

Table 3. Classification accuracies (in percentage) of the H/α method (bold: after compensation).

	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8	Z_9
Z_1	67.6	2.2	0	30.1	0	0	0	0	0
	88.9	0	0	**11**	0	0	0	0	0
Z_2	19.7	65.8	0	9.7	4.6	0	0	0	0
	13.52	**84.2**	0	**1.9**	**0.2**	0	0	0	0
Z_3	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
Z_4	0	0	0	66.2	0	0	33.7	0	0
	0	0	0	**90.1**	0	0	9.9	0	0
Z_5	0	0	0	26.6	69.8	3.5	0	0	0
	0	0	0	**17.9**	**81.3**	0.6	0	0	0
Z_6	0	0	0	0	3.10	94.4	0	0	2.4
	0	0	0	0	0	**100**	0	0	0
Z_7	0	0	0	0	0	0	100	0	0
	0	0	0	0	0	0	**100**	0	0
Z_8	0	0	0	0	0	0	5.9	85.1	8.8
	0	0	0	0	0	0	**0.8**	**96.4**	2.6
Z_9	0	0	0	0	0	0	0	0	100
	0	0	0	0	0	0	0	0	**100**

where n denotes the number of data samples, which corresponds to the number of pixels in our case. u_{ki} is the fuzzy membership of sample x_i to cluster k satisfying

$$\sum_{k=1}^{c} (u_{ki}) = 1, \quad 0 \le u_{ki} \le 1, \quad 0 \le \sum_{i=1}^{n} (u_{ki}) \le 1$$

Parameter $m > 1$ is a weighting factor that controls the fuzziness of the membership function. In our case $m = 1.1$. $d^2(x_i, \nu_k)$ is the distance measure that has many forms depending on the application. In our study, we selected the basic form which is Euclidean distance. By minimizing Eq. (16) using the Lagrange multiplier method, the updated equations of the membership u_{ki} and cluster center ν_k are

$$u_{ki} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d^2 (x_i, \nu_i)}{d^2 (x_i, \nu_k)} \right)^{\frac{1}{m-1}}} \tag{23}$$

$$\nu_k = \frac{\sum_{i=1}^{n} (u_{ki})^m x_i}{\sum_{i=1}^{n} (u_{ki})^m} \tag{24}$$

The FCM algorithm included the following steps:

i. Initialize the cluster centers v_k. The same vector is used to classify original and compensated data.

ii. Update membership values u_{kj} via Eq. (17).

iii. Update cluster centers v_k via Eq. (18).

iv. Repeat steps ii and iii M times. In our case, dataset was introduced 10 times.

- Simulated data.
  Table 4 gives the performances of the FCM using original and compensated data having X_2 as inputs. We observe that the majority of targets were discerned perfectly even by using the original

Table 4. FCM classification accuracies (in percentage) of simulated data using X_2 (bold: after compensation).

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8
T_1	96.4	3.3	0.3	0	0	0	0	0
	98.2	**1.8**	0	0	0	0	0	0
T_2	7.1	92.6	0.2	0	0	0	0	0
	5.5	**94.2**	0.2	0	0	0	0	0
T_3	0	0	99.1	0	0	0.4	0	0.5
	0	0	**100**	0	0	0	0	0
T_4	0	0	0	100	0	0	0	0
	0	0	0	**100**	0	0	0	0
T_5	0	0	0	0	100	0	0	0
	0	0	0	0	**100**	0	0	0
T_6	0	0	0.5	0	0	99.3	0	0.2
	0	0	0	0	0	**100**	0	0
T_7	0	0	0	0	0	0.7	0	99.3
	0	0	0	0	0	0	0	**100**

data. For the remaining targets, we observe that by using Yahia/Aguili's method the performances of the classification increased slightly. Table 5 shows the performances of the FCM using original and compensated data (using AQ-MLE and Yahia/Aguili's method) having X_3 as feature vector. We observe again that Yahia/Aguili's method clearly increased the performances of the classification whereas the AQ-MLE degraded some classification accuracies.

- Real data.
  In this subsection, we classify real data (i.e., zones A, B, C, D and E). Table 6 gives the results of the FCM using original and compensated data using X_2 as input vector. As for simulated data, we observe that by using Yahia/Aguili's bias compensation method, classification accuracies were increased even using the *span* as input parameter. Table 7 displays the performances of the FCM using original and compensated data having X_3 as feature vector. We observe again that Yahia/Aguili's method increased the performances of the classification whereas the AQ-MLE degraded the classification accuracy of some classes.

6.2. Supervised Classification

6.2.1. Multilayer Perceptron Neural Network

MLP is a feedforward artificial neural network trained using a supervised training algorithm such as back-propagation [25]. The MLP structure is described in Fig. 7. Our MLP was three layered. The number of nodes of the input layer is equal to the length of the input feature vector. The number of nodes in the output layer is equal to the number of classes. The number of nodes of the hidden-layer can be estimated using [25]:

$$N_h = \sqrt{N_{in} \times N_o} \tag{25}$$

The MLP training includes the following steps:

i. Randomly set the values of weight matrices WE and WS.

ii. Randomly choose an input vector $X_k = \{x_{1k}, x_{2k}, \ldots, x_{Nink}\}$ from the dataset and set its desired ouput Y^k. k is the kth iteration.

iii. Compute the inputs and the outputs of nodes in Fig. 7 using:

$$IC^k(j) = \sum_{i=1}^{Nin} x_{ik} WE_{(j,i)}^{(k)} \tag{26}$$

$$OC^k(j) = \tanh\left(IC^k(j)\right) \tag{27}$$

$$IS^k(l) = \sum_{j=1}^{No} OC^k(j)WS^{(k)}_{(l,j)} \tag{28}$$

$$S^k(j) = \tanh\left(IS^k(j)\right) \tag{29}$$

Table 5. FCM classification accuracies (in percentage) of simulated data using X_3 (bold: Yahia/Aguili, italic: AQ-MLE).

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8
T_1	73.6	10.8	0	14.6	0	0	0	0.8
	88.6	**9**	0	**1.7**	0	0	0	**0.4**
	76	*12.2*	0	*10.6*	0	0	0	*1.1*
T_2	39.2	57.4	0	3.3	0	0	0	0
	24.8	**74.7**	0	**0.4**	0	0	0	**0.2**
	39.9	*58.3*	0	*1.7*	0	0	0	0
T_3	0	0	95.12	0	0.4	3.3	0.2	0.8
	0	0	**99.1**	0	0	**0.6**	0	0
	0	0	*96.2*	0	0	*2.6*	*0.2*	*0.8*
T_4	15.2	0.4	0.8	77.8	5.5	0	0	0
	0	0	**1.3**	**97.7**	**0.8**	0	0	0
	15	*0.4*	*1.7*	*81.9*	*0.8*	0	0	0
T_5	0	0	1.7	2.6	95.5	0	0	0
	0	0	**0.2**	0	**99.8**	0	0	0
	0	0	*0.2*	*3.5*	*95.3*	0	0	0
T_6	0	0	0.2	0	0	99.7	0	0
	0	0	0	0	0	**100**	0	0
	0	0	0.2	0	0	*99.7*	0	0
T_7	0	0	0	0	0	0.2	79.3	20.3
	0	0	0	0	0	0	**99.7**	**0.2**
	0	0	0	0	0	0	*77.8*	*22.1*
T_8	0	0	0	0	0	2.4	43	54.5
	0	0	0	0	0	**2.4**	**1.7**	**95.7**
	0	0	0	0	0	*1.1*	*47.2*	*51.6*

Table 6. FCM classification accuracies (in percentage) of real data using X_2 (bold: after compensation).

	A	B	C	D	E
A	58.6	0	12.2	2.5	26.7
	63.6	0	**10.4**	**1.5**	**24.5**
B	0	97.3	0	2.7	0
	0	**97.6**	0	**2.4**	0
C	10.1	0	61	0	28.9
	8.1	0	**62.1**	0	**28**
D	2.7	0.6	0	96.2	0.5
	1.8	**1.1**	0	**96.9**	**0.2**
E	14.6	0	23.2	0.2	62
	13.9	0	**21.1**	0	**65**

Table 7. FCM classification accuracies (in percentage) of real data using X_3 (bold: Yahia/Aguili method, italic: AQ-MLE).

	A	B	C	D	E
	77.6	0	0	0	22.4
A	**87.1**	0	0	0	**12.9**
	72.8	*0*	*0*	*0*	*27.2*
	0	96	0	4	0
B	0	**97**	0	**3**	0
	0	*95.3*	*0*	*4.7*	*0*
	16.5	0	81.5	2	0
C	**11**	0	**88.8**	**0.2**	0
	19.7	*0*	*79.7*	*0.6*	*0*
	0.2	0	9.6	90.2	0
D	**0.3**	0	**5.4**	**94.3**	0
	0.7	*0*	*7.2*	*92.1*	*0*
	0	0	0	0	100
E	0	0	0	0	**100**
	0	*0*	*0*	*0*	*100*

Table 8. MLP classification accuracies (in percentage) of simulated data using X_2 (bold: after compensation).

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8
T_1	100	0	0	0	0	0	0	0
	100	0	0	0	0	0	0	0
T_2	0	100	0	0	0	0	0	0
	0	**100**	0	0	0	0	0	0
T_3	0	0	98.9	0.4	0.3	0	0	0.4
	0	0	**100**			0	0	
T_4	0	0	0	100	0	0	0	0
	0	0	0	**100**	0	0	0	0
T_5	0	0	0	0	100	0	0	0
	0	0	0	0	**100**	0	0	0
T_6	0	0	0.7	0	0	99.1	0	0.2
	0	0		0	0	**100**	0	
T_7	0	0	0	0	0	0	100	0
	0	0	0	0	0	0	**100**	0
T_8	0	0	0	0	0	0.4	0	99.6
	0	0	0	0	0		0	**100**

iv. Compute the error $e = \|S^k - Y^k\|$ and update WE and WS using the back-propagation training algorithm [26]. $\|\cdot\|$ is the Euclidean distance.

v. Repeat steps iii and iv L times. In our case, the dataset was introduced 1000 times.

- Simulated data.
  Table 8 shows the performances of the MLP using original and compensated data with X_2 as inputs. As expected, the classification accuracies were noticeably superior to those obtained using unsupervised classification. We observe that by using Yahia/Aguili's procedure, all targets were perfectly discerned. Since the MLP produced high accuracy levels, the impacts of bias
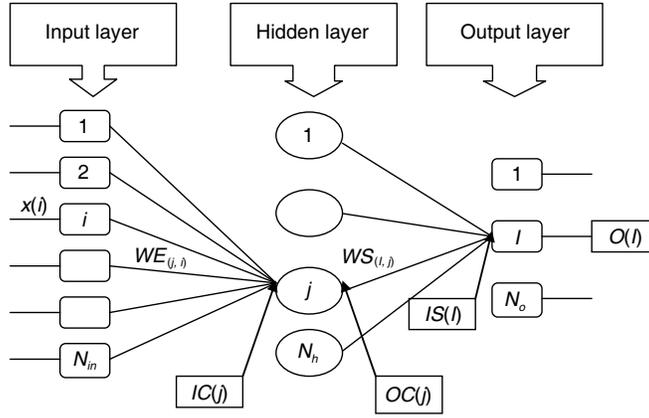
Figure 7. Architecture of the multi-layer perceptron.

Table 9. FCM classification accuracies (in percentage) of real simulated using X_3 (bold: Yahia/Aguili, italic: AQ-MLE).

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8
	99.5	0.5	0	0	0	0	0	0
T_1	**100**	0	0	0	0	0	0	0
	98.7	0.8	0	0	0	0	0	0.4
	2.7	97.3	0	0	0	0	0	0
T_2	**0.2**	**99.8**	0	0	0	0	0	0
	6	*93.5*	0	0.2	0	0	0	0.2
	0	0	96.7	2	0	0.4	0	0.9
T_3	0	0	**99.8**	0	0	0	0	**0.2**
	0	0	*97.4*	*1.6*	*0.2*	*0.2*	0	*0.6*
	0	0	0.2	99.8	0	0	0	0
T_4	0	0	**0.7**	**99.3**	0	0	0	0
	0	0	*0.2*	*99.8*	0	0	0	0
	0	0	0	0	100	0	0	0
T_5	0	0	0	**0.2**	**99.8**	0	0	0
	0.2	0	0	0	*99.8*	0	0	0
	0	0	1.5	0	0	97.8	0	0.7
T_6	0	0	0	0	0	**100**	0	0
	0	0	*2.2*	0	0	*97.8*	0	0
	0	0	0	0	0	0	100	0
T_7	0	0	0	0	0	0	**100**	0
	0	0	0	0	0	0	*100*	0
	0.5	0	0	0	0	3.1	0	96.4
T_8	0	0	0	0	0	**0.4**	0	**99.6**
	0	0	0	0	0	*2.4*	*0.2*	*97.1*

compensation were not very noticeable. Table 9 gives the performances of the MLP using original and compensated data having X_3 as feature vector. We observe again that Yahia/Aguili's method increased the performances of the classification whereas the by using AQ-MLE some classification accuracies were degraded.

- Real data.

Table 10 shows the performances of the MLP using original and compensated data with X_2 as inputs. As for simulated data, we observed that by using Yahia/Aguili's bias compensation method, all classification accuracies increased even when using the *span* as input parameter. However, due to high performances obtained using the MLP, the improvements were not very noticeable. Tables 11 represents the performances of the MLP using original and compensated data using X_3 as feature vector. We observe again that Yahia/Aguili's method increased the performances of the classification whereas the AQ-MLE degraded the classification accuracies.

Supervised and unsupervised classifications obtained using simulated and real data show that Yahia/Aguili's bias compensation method increased the classification accuracies whereas the AQ-MLE method degraded some classification results. As a consequence, to improve the classification accuracies, bias compensation procedures must reduce the speckle noise.

By using feature vector X_2, the classification accuracies obtained after bias compensation did not increase significantly as when using X_1 and X_3. Consequently, when span is added to the biased parameters $(H/\alpha/A)$, the classification performance is improved. Thus, in these cases, bias elimination is not so critical but advantageous.

Table 10. MLP classification accuracies (in percentage) of real data using X_2 (bold: after compensation).

	A	B	C	D	E
A	83.4	0	16.1	0	0.5
	86.4	0	**13.1**	0	**0.5**
B	0	100	0	0	0
	0	**100**	0	0	0
C	15.5	0	84	0	0
	14	0	**86**	0	0
D	0	0	0	100	0
	0	0	0	**100**	0
E	0.3	0	0	0	99.7
	0	0	0	0	**100**

Table 11. MLP classification accuracies (in percentage) of real data using X_3 (bold: Yahia/Aguili, italic: AQ-MLE).

	A	B	C	D	E
A	96.5	0	2.1	0	1.4
	97.5	0	**1.1**	0	**1.4**
	95	0	*1.7*	0	*3.3*
B	0	100	0	0	0
	0	**100**	0	0	0
	0	*99*	0	*1*	0
C	1.4	0	98.6	0	0
	0.7	0	**99.3**	0	0
	5	0	*91.8*	*3.2*	0
D	0	0	0	100	0
	0	0	0	**100**	0
	0.3	*0.4*	0	*99.2*	*0.1*
E	2	0	0	0	98
	0.2	0	0	0	**99.8**
	5	0	0	0	*95*

7. CONCLUSION

In this paper, we studied the impacts of bias compensation on the classification. We considered both supervised and unsupervised classification techniques. In particular, we studied the basic versions of H/α plane, FCM and MLP classifiers because our aim consisted in studying only the effects of bias compensation on the classification. Other more sophisticated versions or other classifiers can be used. We demonstrated that bias elimination was necessary to increase the dynamic ranges of $H/A/\alpha$ parameters, which implies more ability to discern different targets. Nevertheless, bias compensation techniques introduced in the literature compensated the bias but changed the variances of the noise. By applying Yahia/Aguili's method which reduced the noise, both supervised and unsupervised classification accuracies were increased because the overlapping between classes was reduced. Conversely, by applying the AQ-MLE, the classification accuracies were decreased because the noise was amplified. We verified that the combination of the *span* and biased parameters decreased the effects of bias but did not eliminate them totally.

REFERENCES

1. Oliver, C. and S. Quegan, *Understanding Synthetic Aperture Radar Images*, Artech House, Boston, MA, 1998.
2. Cloude, S. and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.*, Vol. 34, No. 2, 498–518, Mar. 1996.
3. Cloude, S. and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sens.*, Vol. 35, No. 1, 68–78, Jan. 1997.
4. Park, S.-E. and W. M. Moon, "Unsupervised classification of scattering mechanisms in polarimetric SAR data using fuzzy logic in entropy and alpha plane," *IEEE Trans. Geosci. Remote Sens.*, Vol. 45, No. 8, 2652–2664, Aug. 2007.
5. Salehi, M., M. R. Sahebi, and Y. Maghsoudi, "Improving the accuracy of urban land cover classification using RADARSAT-2 PolSAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, Vol. 7, No. 4, 1394–1401, Apr. 2014.
6. Bhattacharya, A. and R. Touzi, "Polarimetric SAR urban classification using the Touzi target scattering decomposition," *Can. J. Remote Sens.*, Vol. 37, No. 4, 323–332, Aug. 2011.
7. Antropov, O., Y. Rauste, H. Astola, T. Häme, and M. T. Hallikainen, "Land cover and soil type mapping from spaceborne PolSAR data at L-band with probabilistic neural network," *IEEE Trans. Geosci. Remote Sens.*, Vol. 52, No. 9, 5256–5270, Sep. 2014.
8. Lee, J. S., M. R. Grunes, T. L. Ainsworth, L. Du, D. L. Schuler, and S. R. Cloude, "Unsupervised classification of polarimetric SAR imagery based on target decomposition and Wishart distribution," *IEEE Trans. Geosci. Remote Sens.*, Vol. 37, No. 5, 2249–2258, Sep. 1999.
9. Ferro-Famil, L., E. Pottier, and J. S. Lee, "Unsupervised classification of multifrequency and fully polarimetric SAR images based on the H/A/Alpha — Wishart classifier," *IEEE Trans. Geosci. Remote Sens.*, Vol. 39, No. 11, 2332–2342, Nov. 2001.
10. Yahia, M. and Z. Belhadj, "Unsupervised classification of polarimetric SAR images using neural networks," *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 203–205, Jul. 2003.
11. Dabboor, M., J. M. Collins, V. Karathanassi, and A. Braun, "An unsupervised classification approach for polarimetric SAR data based on the Chernoff distance for complex Wishart distribution," *IEEE Trans. Geosci. Remote Sens.*, Vol. 51, No. 7, 4200–4213, Jul. 2013.
12. Cao, F., W. Hong, Y. Wu, and E. Pottier, "An unsupervised segmentation with an adaptive number of clusters using the SPAN/$H/\alpha/A$ space and the complex Wishart clustering for fully polarimetric SAR data analysis," *IEEE Trans. Geosci. Remote Sens.*, Vol. 45, No. 11, 3454–3467, Nov. 2007.
13. Zhang, Y., L. Wu, and G. Wei, "A new classifier for polarimetric SAR images," *Progress In Electromagnetics Research*, Vol. 94, 83–104, 2009.
14. Gou, S., X. Qiao, X. Zhang, W. Wang, and F. Du, "Eigenvalue analysis-based approach for Pol-SAR image classification," *IEEE Trans. Geosci. Remote Sens.*, Vol. 52, No. 2, 805–818, Feb. 2014.

15. Goodman, J. W., "Some fundamental properties of speckle," *J. Opt. Soc. Amer.*, Vol. 66, No. 11, 1145–1150, Nov. 1976.

16. Lee, J. S., "Speckle analysis and smoothing of synthetic aperture radar images," *Comput. Graph. Image Process.*, Vol. 17, 24–32, 1981.

17. Yahia, M. and Z. Belhadj, "Polarimetric SAR denoising using adaptive prediction technique," *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 4025–4027, Jul. 2006.

18. López-Martínez, C., E. Pottier, and S. Cloude, "Statistical assessment of eigenvector-based target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.*, Vol. 43, No. 9, 2058–2074, Sep. 2005.

19. Foucher, S., G. Farage, and G. B. Benie, "Application of bootstrap techniques for the estimation of target decomposition parameters in radar polarimetry," *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2224–2228, Jun. 2007.

20. Lee, J. S., T. Ainsworth, J. Kelly, and C. López-Martínez, "Evaluation and bias removal of multi-look effect on entropy/alpha/anisotropy in polarimetric SAR decomposition," *IEEE Trans. Geosci. Remote Sens.*, Vol. 46, No. 10, 3039–3052, Oct. 2008.

21. López-Martínez, C., A. Alonso-González, and X. Fàbregas, "Perturbation analysis of eigenvector-based target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.*, Vol. 42, No. 4, 2081–2086, Apr. 2014.

22. Yahia, M. and T. Aguili, "Characterization and correction of multilook effects on eigendecomposition parameters in PolSAR images," *IEEE Trans. Geosci. Remote Sens.*, Vol. 53, No. 9, 5237–5246, Sep. 2015.

23. Lee, J. S., M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sens.*, Vol. 37, No. 5, 2363–2373, Sep. 1999.

24. Bezdek, J. C., R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computer and Geosciences*, Vol. 10, 191–203, 1984.

25. Duquenoy, M., J. P. Ovarlez, L. Ferro-Famil, and E. Pottier, "Supervised classification of scatterers on SAR imaging based on incoherent polarimetric time-frequency signatures," *European Signal Processing Conference*, 764–768, Aug. 2009.

26. Lippmann, R. P., "An introduction to computing with neural nets," *IEEE ASSP Magazine*, 4–22, Apr. 1994.