MODEL-BASED STATISTICAL ANALYSIS OF POLSAR DATA.

Torbjørn Eltoft[1,2], *Anthony Doulgeris*[1] *and Stian N. Anfinsen*[1]

[1] Department of Physics & Technology, University of Tromsø
[2] Northern Research Institute (Norut), Tromsø,
Email: torbjorn.eltoft, anthony.doulgeris, stian.normann.anfinsen all at @uit.no

ABSTRACT

In this paper, we consider statistical analysis of PolSAR data in the framework of the multivariate product model. The complex scattering vector is here considered as a double stochastic circular Gaussian variable, in which the variance is linearly scaled by a common stochastic scaling factor z. The scaling factor is associated with texture. We discuss various parametric probability density functions for z, and indicate how model parameters can be estimated from data by a simple moment based method. Experimental analysis shows that for some surface covers, certain texture distributions fit better than others. Then, polarimetric covariance matrix data analysis is addressed in the framework of product models, and we propose a processing scheme which perform image segmentation using a stochastic EM approach.

I. INTRODUTRODUCTION

Polarimetric SAR (PolSAR) data are complex multidimensional image data, which can be analyzed along several processing schemes. In the literature, we find that much emphasis has been put on analysis based on target decomposition theorems [1], [2]. Through this approach, information about scattering mechanisms can be gained. The knowledge of the exact statistical properties of PolSAR data founds the basis for another strategy of multidimensional image analysis, which in some cases is complementary to the target decomposition approach. Statistical properties can be used to discriminate different types of land cover (e.g. [3], [4]), or to device specialized filters for speckle noise reduction (e.g. [5]), to mention two areas of application.

It is well known that SAR (and PolSAR) data can be non-Gaussian in nature. For this reason, various non-Gaussian models have been proposed to represent SAR data. These have later been extended into the polarimetric realm, where the multivariate K-distributions [6] and G-distributions [7] are successful examples. Both these distributions are members of the so-called *product model*, which states that, under certain conditions, the backscattered signal results from the product between a Gaussian speckle noise component and the terrain backscatter. Several distributions could be used for the terrain backscatter, in order to model different types of surface classes with their characteristic spatial correlation properties and degrees of homogeneity.

Mathematically, the multivariate product model can be formulated as follows: Let $\mathbf{x}$ be a d-dimensional, zero mean Gaussian variable with covariance matrix equal to the identity matrix. Let furthermore, $\mathbf{\Gamma} \in \mathcal{R}^{d\times d}$ be a positive definite, Hermitian matrix with determinant $\det \mathbf{\Gamma} = 1$, and let Z be a scalar random variable with pdf $f_z(z)$, which can attain only positive values. A new variable $\mathbf{y}$ is now generated as

$$\boldsymbol{y} = \sqrt{z}\mathbf{\Gamma}^{\frac{1}{2}}\boldsymbol{x}. \qquad (1)$$

The matrix $\mathbf{\Gamma}$ defines the internal covariance structure of the component variables of $\mathbf{y}$. For this reason we will refer to this matrix as the covariance structure matrix.

The above modeling scheme constitutes flexible models, which have the capabilities to model data which ranges in Gaussianity from highly non-Gaussian to Gaussian data. The rich parameter space associated with the product models allows for the development of image segmentation algorithms, both in data space and in the parametric feature space.

In this paper some characteristics related to product models are briefly discussed, as well as certain approaches to the analysis of PolSAR data in this framework. Experimental results from various application areas will be presented.

II. PARAMETRIC MODELS FOR VECTOR OBSERVATIONS

The product model scheme describes different parametric families of distributions, depending on the scale parameter probability density function, $f_z(z)$. Given the pdf for the scale parameter, the marginal pdf for $\mathbf{y}$ can be obtained by integrating the conditional pdf of $\mathbf{y}|z$, which is multivariate Gaussian, over the density of z. That is

$$f_{\mathbf{y}}(\mathbf{y}) = \int_0^\infty f_{\mathbf{y}|z}(\mathbf{y}|z)\ f_z(z)\ dz \qquad (2)$$

As can be seen, the probability distribution of the resulting double stochastic variable is obtained as a continuous mixture of scaled Gaussians. For this reason we often refer to these models as scale mixture of Gaussian (SMoG) models, or normal variance mixture variates.

Table I. Scale mixture of Gaussian models

z distribution	Multivariate scale mixture distribution
Constant $\delta(z-\sigma^2)$	Gaussian, $\mathbf{MG}(\mathbf{y};\sigma^2,\boldsymbol{\mu},\boldsymbol{\Gamma})$
Exponential $(z;\lambda) =$ $\frac{1}{\lambda}\exp(-\frac{z}{\lambda})$	Laplacian, $\mathbf{ML}(\mathbf{y};\lambda,\boldsymbol{\mu},\boldsymbol{\Gamma}) =$ $\frac{1}{\pi^d}\,\frac{2}{\lambda}\,\frac{K_{d-1}\left(2\sqrt{\frac{q(\mathbf{y})}{\lambda}}\right)}{\left(\sqrt{\lambda q(\mathbf{y})}\right)^{d-1}}$
Gamma $(z;\alpha,\mu_z) =$ $\left(\frac{\alpha}{\mu_z}\right)^{\alpha}\frac{z^{\alpha-1}}{\Gamma(\alpha)}\,\exp\left(-\frac{\alpha}{\mu_z}z\right)$	K-distribution, $\mathbf{MK}(\mathbf{y};\alpha,\mu_z,\boldsymbol{\mu},\boldsymbol{\Gamma}) =$ $\frac{2}{\pi^d\,\Gamma(\alpha)}\left(\frac{\alpha}{\mu_z}\right)^{\frac{\alpha+d}{2}}(q(\mathbf{y}))^{\frac{\alpha-d}{2}}K_{\alpha-d}\left(2\sqrt{\frac{\alpha\,q(\mathbf{y})}{\mu_z}}\right)$
Inverse Gaussian $(z;\delta,\gamma) =$ $\frac{\delta}{\sqrt{2\pi}}e^{\delta\gamma}\,z^{-\frac{3}{2}}\,\exp\left(-\frac{1}{2}(\frac{\delta^2}{z}+\gamma^2 z)\right)$	Normal Inverse Gaussian, $\mathbf{MNIG}(\mathbf{y};\delta,\gamma,\boldsymbol{\mu},\boldsymbol{\Gamma}) =$ $\sqrt{2}\,\delta e^{\delta\gamma}\left(\frac{\gamma}{\pi\sqrt{\delta^2+2q(\mathbf{y})}}\right)^{d+\frac{1}{2}}K_{d+\frac{1}{2}}\left(\gamma\sqrt{\delta^2+2q(\mathbf{y})}\right)$

$q(\mathbf{y}) = (\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Gamma}^{-1}(\mathbf{y}-\boldsymbol{\mu})$ is the scaled squared Mahalonobis distance from the mean, with $\boldsymbol{\mu}=\mathbf{0}$ for the PolSAR case

$K_m(x)$ is a modified Bessel function of the second kind with order m

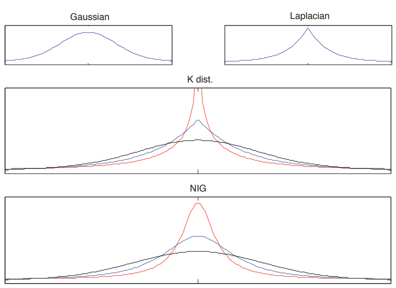


Fig. 1. Example shapes for each model distribution, fixed width. The two parameter models, the K-distribution and the normal inverse Gaussian, can vary in shape.

Four scale mixture models, derived with closed form expressions in [8], are depicted in Table I, including the Multivariate Gaussian (MG) distribution as a special case. All are sparse (heavy tailed) and symmetric distributions, with a global shape for all dimensions, but an allowable width variation described by the covariance structure matrix (see Fig.1). Both the MG and multivariate Laplacian (ML) distributions have fixed shapes, and only vary with width parameter. The two-parameter Multivariate K-distribution (MK) and multivariate normal inverse Gaussian (MNIG) distributions describe a range of shapes as well as widths, both including the MG as a limiting shape. Both the MK and the MNIG distribution have theoretical links to the nature of distributed target scattering, as they may be viewed as statistical models associated with Brownian motion [9].

Given such a general scheme as in (1), it can be readily shown that:

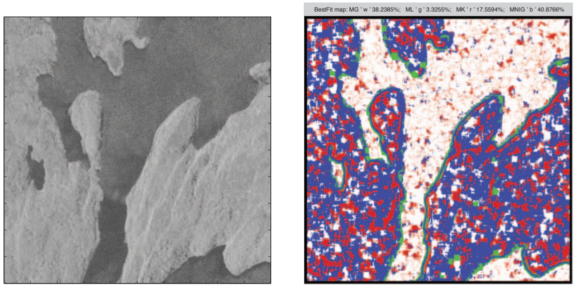
$$\mathbf{E}\{\mathbf{y}\} = \boldsymbol{\mu} \quad (=\mathbf{0}) \qquad (3)$$

$$\mathbf{E}\{(\mathbf{y}-\boldsymbol{\mu})(\mathbf{y}-\boldsymbol{\mu})^H\} = \mathbf{E}\{z\}\boldsymbol{\Gamma} \quad (=\boldsymbol{\Sigma}) \qquad (4)$$

$$\mathbf{E}\{[(\mathbf{y}-\boldsymbol{\mu})^H\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})]^2\} = \frac{\mathbf{E}\{z^2\}}{[\mathbf{E}\{z\}]^2}d(d+1) \qquad (5)$$

where $(\cdot)^H$ means (Hermitian) conjugate transpose, $\mathbf{E}\{\cdot\}$ is the expectation operator, and d is the dimension (which will be 4 for PolSAR data). These equations can be used to estimate the various parameters: obtaining $\boldsymbol{\Gamma}$ and $\mathbf{E}\{z\}$ from the sample covariance via (4) plus the normalisation $|\boldsymbol{\Gamma}| = 1$, and the second moment $\mathbf{E}\{z^2\}$ from the sample multivariate kurtosis via (5) [10]. It is mathematically convenient to define a scale invariant measure of non-Gaussianity, the relative kurtosis (RK) as $\frac{\mathbf{E}\{z^2\}}{[\mathbf{E}\{z\}]^2}$, which is easily found from (5). The particular parametric form of the distribution of z is then obtained by the solution for its first and second moments given the estimates obtained for $\bar{z} = \mathbf{E}\{z\}$ and RK from (4) and (5).

In the general case, all the model parameters are free to be optimised in the fitting procedures. However, if we have some *a priori* knowledge about a parameter's value, we would expect a better model fit by actually constraining it. The most obvious constraint in our radar data is the expected zero-mean. And, if for example reflection symmetry can be assumed valid, we have further zero and pair value constraints on the covariance structure matrix. Simulated studies show that applying the constraints improves the parameter estimation greatly, particularly when the sample size is small, with the covariance constraints being the most significant.



(a) Mountain lake and forest

Fig. 2. Intensity and 'best fit' coloured map. Gaussian in white, Laplacian in green, K-dist. in red and normal inverse Gaussian in blue.

II-A. Parameter feature space analysis

Statistical modelling can now be achieved by a local neighborhood approach, fitting the parametric models to the collection of data points around each image location by a fast moment-based method. We have used a $13 \times 13 = 169$ window size. Several different real PolSAR data images were analyzed to compare the behaviour of the models for quite different terrain types: (a) mountain lake and forest area ('Bleikvatnet', Norway, EMISAR), (b) a mountain glacier area ('Okstinden', Norway, EMISAR), (c) a sea ice image from (Canada, CONVAIR), and (d) an agricultural and urban area ('Foulum', Denmark, EMISAR). To investigate the suitability of scale mixture models, we utilised the log-likelihood function to test the goodness-of-fit of each of our four models, from which we can visualise a best fit image, colour coded to where each of the four models had the highest goodness-of-fit score. Fig.2 shows the "bets fit image corresponding to the 'Bleikvatnet' data set. From our analysis the following observations may be made:

- Uniform, smooth or homogeneous areas are usually best fitted as Gaussian.
- The land in general, the visible icy crevasses, rocky outcrops, urban areas and certainly anything with small scale details and high contrast are certainly non-Gaussian in nature and were poorly fitted by the Gaussian model.
- All types of vegetated land appear to be best described by the normal inverse Gaussian distribution, whereas the sea ice image by the K-distribution, although the difference compared to the NIG was negligible.
- The urban areas and coastlines best fitted more often by the Laplacian, however this may be due to high contrast edge mixture effects because it appears at all water/land boundaries, around point sources like known huts within the forest, and along hedge/fence lines around fields.

III. PARAMETRIC MODELS FOR MATRIX OBSERVATIONS

PolSAR data are often available as multi-looked matrix data. From the observation vector $\boldsymbol{y}$ we form the covariance matrix $\boldsymbol{C}$ as

$$\boldsymbol{C} = \frac{1}{L}\sum_{i=1}^{L} \boldsymbol{y}\boldsymbol{y}^H \tag{6}$$

If the vector is assumed to be complex multivariate Gaussian, then the complex sample covariance matrix formed by a multi-look average over L individual vector points will be complex Wishart distributed. We denote this as $\mathbf{C} \sim \mathcal{W}_d(\boldsymbol{C}; L, \boldsymbol{\Sigma})$ with L degrees of freedom, centered around the true covariance. The SMoG models may be applied to generate non-Wishart distributions for matrix variables. The explicit form is again determined by the parametric distribution for the texture variable z. Table II summarizes two models previously presented in the litterature [6], [7].

Estimation in the case of L-look data is based upon the neighbourhood mean of the matrix-variate data, plus the variance of a mean squared Mahalanobis measure (M) which is equivalent to trace($\boldsymbol{\Sigma}^{-1}C$). Assuming that $\boldsymbol{\mu} = \mathbf{0}$ for simplicity, it is easily shown that:

$$\begin{aligned}
\mathbf{C} &= \frac{1}{L}\sum_{l=1}^{L} \mathbf{y}\mathbf{y}^H && (7)\\
\mathbf{E}\{\mathbf{C}\} &= \mathbf{E}\{z\}\boldsymbol{\Gamma} \quad (= \boldsymbol{\Sigma}) && (8)\\
M &= \text{tr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{C}\right) && (9)\\
&= \frac{1}{L}\sum_{l=1}^{L} \left(\mathbf{y}_l^H \boldsymbol{\Sigma}^{-1}\mathbf{y}_l\right) && (10)\\
\mathbf{E}\{M\} &= d && (11)\\
\mathbf{E}\{(M-d)^2\} &= \frac{1}{L}\frac{\mathbf{E}\{z^2\}}{[\mathbf{E}\{z\}]^2} d(d+1) \quad - \quad \frac{1}{L}d^2 && (12)
\end{aligned}$$

Note that the expectation of M equals d because of the normalisation with respect to each local covariance matrix $\boldsymbol{\Sigma}$ in tr($\boldsymbol{\Sigma}^{-1}\boldsymbol{C}$). The parameters $\boldsymbol{\Gamma}$ and $\mathbf{E}\{z\}$ are obtained from the mean matrix by applying the constraint that $|\boldsymbol{\Gamma}| = 1$, and RK is obtained in terms of var(M) by rearranging (12).

III-A. Unsupervised classification

We have studied an EM style algorithm to perform unsupervised classification based on the K-Wishart probability distribution. That is, we assign a pixel to the class $\omega_i, i \in 1, ..., k$ that is most probable given the data sample $\boldsymbol{C}$. The number of classes k must be determined in advance. We thus select the class that maximises

$$P(\omega_i|C;\boldsymbol{\theta}_i) \propto f_C(C|\omega_i;\theta_i)P(\omega_i) \tag{13}$$

Here $\boldsymbol{\theta}_i = \{\alpha_i, \mu_i, \boldsymbol{\Gamma}_i\}$ is defined as the parameters of the K-Wishart distribution modelling class ω_i. Starting with some initial class estimates, we re-classify each pixel to the class maximising (13), re-calculate the model parameters

Table II. Covariance matrix distributions for two texture models

$f_z(\tau)$ of texture variable z		$f_{\mathbf{C}}(\mathbf{C})$ of covariance matrix $\mathbf{C}$	
Constant	$\delta(z-\sigma^2)$	$\mathcal{W}_d(\boldsymbol{C};\boldsymbol{\Sigma},L)$	$\frac{L^{Ld}}{\Gamma_d(L)}\frac{\lvert\mathbf{C}\rvert^{L-d}}{\lvert\boldsymbol{\Sigma}\rvert^{L}}\exp\left\{\mathrm{tr}(-L\boldsymbol{\Sigma}^{-1}\mathbf{C})\right\}$
Gamma $(z;\alpha,\mu)$	$\left(\frac{\alpha}{\mu_z}\right)^{\alpha}\frac{z^{\alpha-1}}{\Gamma(\alpha)}\exp\left(-\frac{\alpha}{\mu_z}z\right)$	$\mathcal{K}_d(\boldsymbol{C};\boldsymbol{\Sigma},L,\alpha,\mu_z)$	$\frac{2\lvert\mathbf{C}\rvert^{L-d}}{I(L,d)\Gamma(\alpha)}\left(\frac{L\alpha}{\mu_z}\right)^{\frac{\alpha+Ld}{2}}\left(\mathrm{tr}(\boldsymbol{\Gamma}^{-1}\mathbf{C})\right)^{\frac{\alpha-Ld}{2}}K_{\alpha-Ld}\left(2\sqrt{\frac{L\alpha}{\mu_z}\mathrm{tr}(\boldsymbol{\Gamma}^{-1}\mathbf{C})}\right)$
Inverse Gamma $\bar{\gamma}^{-1}(\lambda)$	$\frac{(\lambda-1)^{\lambda}}{\Gamma(\lambda)}\frac{1}{z^{\lambda+1}}\exp\left(-\frac{\lambda-1}{z}\right)$	$\mathcal{G}_d^0(\boldsymbol{C};\boldsymbol{\Sigma},L,\lambda)$	$\frac{L^{Ld}\lvert\mathbf{C}\rvert^{L-d}}{\lvert\boldsymbol{\Sigma}\rvert^{L}}\frac{\Gamma(Ld+\lambda)(\lambda-1)^{\lambda}}{\Gamma_d(L)\Gamma(\lambda)}\left(L\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{C})+\lambda-1\right)^{-\lambda-Ld}$

based on the new partitioning, and repeat these two steps iteratively until convergence (defined by some stop criterion). The resulting class partition is then the classication of the given pixels, under the k mixtures with parameters, $\boldsymbol{\theta}_i$ for $i = 1...k$. Fig.3 illustrates the result of applying this

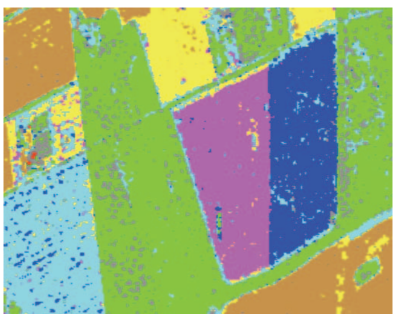


Fig. 3. K-Wishart clustering (Foulum data set).

procedure to a small subset of the Foulum EMISAR data set. More details on theoretical issues and experimental results using this method can be found in [6].

IV. CONCLUSION

We have presented the general framework for generating non-Gaussian statistical distributions known in the literature as the product model (or scale mixture of Gaussian model). In the context of PolSAR data studies this framework offers new strategies in which these data can be analyzed. Statistical modelling is achieved by a local neighborhood approach, fitting the parametric models to the collection of data points around each image location by a fast moment-based method. In this way, PolSAR images are transformed into feature images which indicate where different features are significant. It also turns out that the parametric feature set derived from the modelling has a rich distinguishing power, and the model parameters hence can be used to perform image segmentation. It is also indicated how polarimetric covariance matrix data analysis may be addressed in the framework of product models, and we propose a processing scheme which perform image segmentation using a stochastic EM approach.

V. REFERENCES

[1] S. Cloude and E. Pottier, " A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sensing.*, vol. 34, no. 2, pp. 498-518, 1996.

[2] T. Freeman, and S.L. Durden, "A Three-Component Scattering Model for Polarimetric SAR Data," *IEEE Trans. Geosci. Remote Sensing.*, vol. 36, no. 3, pp.963-973, 19968.

[3] S. H. Yueh, J. A. Kong, J. K. Jao, R. T. Shin, and L. M. Novak, "K-Distribution and polarimetric terrain radar clutter", *J. Electro. Waves Applic.*, vol., pp. 747-768, 1989.

[4] J. Lee, M. Grunes, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on the complex Wishart distribution," *Int. J. Remote Sensing*, vol. 15, no. 11, 1994.

[5] T. Eltoft, "Modeling the amplitude statistics of ultrasonic images," *IEEE Trans. Med. Imag.*, vol. 25, no. 2, 2006.

[6] A. Doulgeris, S. Annsen, and T. Eltoft, Classication with a non-Gaussian model for PolSAR data, *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 10, pp. 2999-3009, 2008.

[7] C. C. Freitas, A. C. Frery, and A. H. Correia, "The polarimetric G distribution for SAR data analysis," *Environmetrics*, vol. 16, pp. 13–31, 2005.

[8] T. Eltoft, T. Kim, and T.-W. Lee, "Multivariate scale mixture of Gaussians models," in *Proceedings of ICA 2006, Charleston, SC, USA*, Mar. 2006.

[9] O. E. Barndorff-Nielsen, "Normal Inverse Gaussian Distributions and Stochastic Volatility Modeling," *Scand. J. Statist.*, vol. 24, pp. 1-13, 1997.

[10] K. V. Mardia, "Measure of Multivariate Skewness and Kurtosis with Applications," *Biometrica*, vol. 57, 3, pp. 519–530, Dec. 1970.