The GA0 Distribution as the True Model for SAR Images

Julio Jacobo-Berlles[1], Marta E. Mejail[1], Alejandro C. Frery[2]

[1]Departamento de Computación, Universidad de Buenos Aires, Ciudad Universitaria
1428 Buenos Aires, Argentina
jacobo,marta@dc.uba.ar
[2]Departamento de Informática, Universidade Federal de Pernambuco, Recife, PE, Brazil
frery@di.ufpe.br

Abstract. For many SAR sensors, urban areas, forests and pasture areas have a growing degree of homogeneity, and for each one of these region types there is a distribution class that fits the data best. In this work, the feasibility of using the $\mathcal{G}_A^0$ distribution as a general model is studied, and the practical usefulness of this proposal is shown. The $\mathcal{G}_A^0$ distribution will be adopted as the true model, and its locally estimated parameters will be used for classification of synthetic and real SAR images.

1 Introduction

Data from a coherent illumination system like a Synthetic Aperture Radar (SAR) can be modelled by the multiplicative model. This model states that, under certain conditions, the constructive and destructive interference of the incident and reflected signals produce a return that varies, in a random manner, with the mean value of the backscatter corresponding to the illuminated target.

This type of data suffers from the presence of a noise called speckle, typical of all images generated with coherent illumination, like ultrasound and laser images. In SAR images, the value of each pixel is a complex number, but in many applications, only the modulus of these complex numbers (amplitude images) or their squared value (intensity images) are used. In this work linear detection (amplitude) data will be used.

The observed value (return) in amplitude format is modelled as a random variable resulting from the product of two independent random variables, which correspond to the backscatter and to the speckle noise. The statistical model for the speckle noise is the Square Root of Gamma, and the model for the backscatter will depend on several parameters that relate to the roughness and texture of the target. The parametric modelling of the backscatter will be treated here, since the estimation of these parameters plays a central role in image analysis.

For regions where the backscatter can be considered constant, like crops and pastures, the $\Gamma^{1/2}$ distribution is a good model for the returned data. The $\mathcal{K}_A$ distribution models the return from homogeneous zones (with certain restrictions due to numerical problems) as well as heterogeneous ones, like forest on flat relief. The $\Gamma^{1/2}$ distribution, however, does not fit heterogeneous data appropriately.

When the area under study is extremely heterogeneous, as is the case for urban areas or forest over undulated relief, both the $\Gamma^{1/2}$ distribution and the $\mathcal{K}_A$ distribution do not model the data adequately. In this case the $\mathcal{G}_A^0$ distribution behaves remarkably well. Taking into account that this distribution models very well data from heterogeneous and homogeneous areas too, and that it is more computationally and theoretically tractable, the substitution of the $\mathcal{G}_A^0$ distribution for the $\mathcal{K}_A$ distribution is proposed, and the feasibility of this substitution is studied. To this end, a correspondence between the parameters of both distributions will be considered in order to approximate, in some sense, the $\mathcal{G}_A^0$ distribution to the KA distribution. This study consists of two parts:

1. Minimisation L_2 in of the distance between the respective densities, in order to obtain a correspondence between distributions.
2. The goodness of fit of the $\mathcal{G}_A^0$ distribution to $\mathcal{K}_A$ distributed data using the test in a Monte Carlo experiment.

A new approach to classification, based on features extraction, for simulated and real SAR images using as true model the $\mathcal{G}_A^0$ distribution, will be shown.

2 Main properties

The $\mathcal{G}_A^0(\alpha,\gamma,n)$ and the $\mathcal{K}_A(\alpha,\lambda,n)$ distributions we will use are characterised by the following densities

$$f_G(z) = \frac{2n^n\Gamma(n-\alpha_G)z^{2n-1}}{\Gamma(n)\Gamma(-\alpha_G)\gamma^{\alpha_G}(\gamma+nz^2)^{n-\alpha_G}}, \qquad (1)$$

where $-\alpha_G, \gamma, n, z > 0$ and

$$f_K(z) = \frac{4(\lambda n)^{(\alpha_K+n)/2}}{\Gamma(n)\Gamma(\alpha_K)} z^{n-\alpha_K+1} K_{n-\alpha_K}(2z\sqrt{\lambda n}), \tag{2}$$

where $\alpha_K, \lambda, n, z > 0$ and K_ν is the modiffied Bessel function of the third kind and order ν, respectively.

There are not many computational implementations of this Bessel function (see [3], for a recent algorithm). On the other hand, the only special function in the $\mathcal{G}_A^0$ distribution is the Γ function, for which there are many reliable implementations.

This is the first computational argument in favour of the $\mathcal{G}_A^0$ distribution and against the $\mathcal{K}_A$ distribution. The second computational argument aims in the same direction and requires the definition of the cumulative distribution functions for both distributions.

Let the random variable V be $\mathcal{G}_A^0(\alpha, \gamma, n)$ distributed. Its cumulative distribution function is given by

$$\Pr(V \leq v) = \frac{n^{n-1}\Gamma(n-\alpha_G)}{\gamma^n \Gamma(n)\Gamma(-\alpha_G)} v^{2n} H(n, n-\alpha_G, n+1; -\frac{n}{\gamma}v^2),$$

where H is is the hypergeometric function. This cumulative distribution function is easy to evaluate using the Snedecor's F distribution, since it can be seen that

$$\Pr(V \leq v) = \Upsilon_{2n,-2\alpha_G}\left(-\frac{\alpha_G}{\gamma} z^2\right), \tag{3}$$

where $\Upsilon_{\phi,\zeta}$ is the cumulative distribution function of a Snedecor's F distributed random variable with ϕ and ζ degrees of freedom.

Consider now $W \sim \mathcal{K}_A(\alpha, \lambda, n)$ In order to write the cumulative distribution function of this random variable it is necessary to impose restrictions on the variation domain of its parameters. Originally, the parametric space of the $\mathcal{K}_A(\alpha, \lambda, n)$ distribution is $(R^+)^3$ but, to be able to write its cumulative distribution function in a recursive form, it is necessary to restrict the variation of either n or α_K to the integers. For the former, this function is given by

$$\Pr(W \leq w) = 1 + \frac{2^{2-\alpha_K-n}}{\Gamma(\alpha_K)\Gamma(n)} g(\nu, k, z), \tag{4}$$

where $z = 2w\sqrt{\alpha n}$, $k = 2n-1$, $\nu = \alpha_K - n$ and the function $g(\nu, k, z) =$ is given by the recursive relation

$$\begin{cases} -z^{\nu+1}K_{\nu+1}(z) & \text{if} \quad k = 1 \\ (k-1)(2\nu+k-1)g(\nu, k-2, z) + & \\ -z^{\nu+k}K_{\nu+1}(z) + & \text{else.} \\ -(k-1)z^{\nu+k-1}K_\nu(z) & \end{cases}$$

More details about this recursive solution for the cumulative distribution function of $\mathcal{K}_A$ distributed variables can be found in [6].

From these considerations we can deduce the following advantages of the $\mathcal{G}_A^0$ distribution over the $\mathcal{K}_A$: it is easier to implement, it uses reliable and immediately obtainable implementations, and it does not impose restrictions on the original parameter space.

The importance of the availability of reliable implementations of the cumulative distribution function arises from the need of carrying out goodness of fit tests and from the use of these functions in estimators based on order statistics.

To the stated advantages, additional advantages in the fields of inference favour even more the use of the $\mathcal{G}_A^0$ distribution instead of the $\mathcal{K}_A$ distribution.

For the estimation of the homogeneity parameter of both distributions using maximum likelihood, the estimator of is α_K difficult to calculate due to the presence of the derivative of the Bessel function with respect to the order parameter. The maximum likelihood estimator of α_G entails the use of the digamma function (Ψ), which has been widely studied and implemented.

Consider the sample $z_1, \ldots, z_k$ of independent observations. The maximum likelihood estimator of α_K, knowing n and λ is

$$k\Psi(\hat{\alpha}_K) - \sum_{i=0}^{k} \frac{\partial}{\partial \hat{\alpha}_K} \log K_{\hat{\alpha}_K - n}(2z_i\sqrt{\lambda n}) = \frac{k}{2}\log\lambda \sum_{i=0}^{k} \log z_i.$$

Analogously, the maximum likelihood estimator of α_G, knowing n and γ, is

$$\Psi(n-\hat{\alpha}_G) - \Psi(\hat{\alpha}_G) = -\log\gamma + \frac{1}{k}\sum_{i=0}^{k} \log(\gamma + n z_i^2). \tag{5}$$

3 Minimisation of the distance between distributions

In this section, a method by which a $\mathcal{G}_A^0$ distribution approximates a $\mathcal{K}_A$ distribution, under the constraint that both must have a mean value equal to one, is described. In other words, given the sets of $\mathcal{G}_A^0$ and $\mathcal{K}_A$ distributions with unitary mean, a correspondence of elements of the first one to elements of the second one, trough a numerical minimisation of an also numerical integration, will be sought. This correspondence will then be established by parameter pairs.

These distributions have only one free parameter: the homogeneity parameter α_K (α_G resp.), because the

scale parameter λ (γ resp.) must be chosen as a function of the number of looks n and the homogeneity parameter, in order to guarantee the constraint.

In order to approximate the $\mathcal{K}_A$ distribution by the $\mathcal{G}_A^0$ distribution it is necessary to previously define the approximation criterion. Consider $\mathcal{D}$, the set of all the distributions that admit a density and establish the notion of proximity between distributions in $\mathcal{D}$ with the distance $d : \mathcal{D} \times \mathcal{D} \rightarrow [0, \infty]$ given by the relation

$$d(D_1, D_2) = \int_{-\infty}^{\infty} |f_1(z) - f_2(z)| \, dz,$$

where f_1 and f_2 are the densities that characterise the distributions D_1 and D_2, respectively. This metric has been already used in a similar context in [4].

We the want to find the value of α_G that minimises the distance $d(\mathcal{G}_A^0(\alpha_G, \gamma, n), \mathcal{K}_A(\alpha_K, \lambda, n))$, using the densities presented in eq. (1) and in eq. (2), where the values $\lambda = \lambda(\alpha_K, n)$ and $\gamma = \gamma(\alpha_G, n)$ are the ones that make the mean value equal to one. Then, the value of α_G that minimises that integral will be sought numerically.

Although α_K varies over all the positive real numbers, for the purposes of this study the search will be done within the interval $[4, 12]$. Very small values of α_K ($0 < \alpha_K < 4$ for instance) correspond to data from extremely heterogeneous areas, which are not well modelled by the $\mathcal{K}_A$ but by the $\mathcal{G}_A^0$ distribution, as can be seen in [2]; then, for these data, it is not necessary to have an approximation. For values of α_K larger than 15, the observed data can be modelled by the $\Gamma^{1/2}$ distribution [2, 6, 7], which is also a particular case of the $\mathcal{G}_A^0$ distribution. So, the only region in which the approximation is necessary is in the interval $[4, 12]$.

For these values of α_K, the corresponding values of α_G obtained as a result of the minimisation of the integral, are shown in Table 1.

4 Goodness of fit of simulated data

To measure the goodness of the fit of $\mathcal{K}_A$ distributed data using the $\mathcal{G}_A^0$ distribution, we will use the χ^2 adherence test in a Monte Carlo experiment. To test if the simulated $\mathcal{K}_A$ distributed data can be fitted by the $\mathcal{G}_A^0$ distribution, the Pearson´s χ^2 statistic will be used:

$$\chi^2 = \sum_{i=0}^{k} \frac{(h_i - mp_i)^2}{mp_i},$$

where m is the total number of $\mathcal{K}_A$ distributed data, h_i is the number of data in each interval, k is the number of intervals, $p_i = F(z_i) - F(z_{i-i})$, where F is the cumulative distribution funtion and the interval i is given by $[z_{i-1}, z_i]$.

α_K	α_G	λ	γ
$n = 1$			
4	-4.3	2.95	4.15
5	-5.3	3.73	5.42
6	-6.3	4.52	6.09
7	-7.3	5.30	7.96
8	-8.3	6.09	9.23
9	-9.3	6.87	10.50
10	-10.3	7.66	11.78
$n = 2$			
4	-4.4	3.32	3.68
5	-5.4	4.20	4.81
6	-6.4	5.08	5.94
7	-7.4	5.96	7.07
8	-8.3	6.85	8.21
9	-9.3	7.73	9.34
10	-10.3	8.61	10.47
$n = 4$			
4	-2.9	3.53	3.45
5	-5.5	4.46	4.53
6	-6.4	5.40	5.59
7	-7.4	6.34	6.65
8	-8.4	7.28	7.72
9	-9.4	8.22	8.78
10	-10.4	9.16	9.84

Table 1: Correspondence between parameters for different numbers of looks.

Let $\mathbf{z_A}$ be a sample of size m of $\mathcal{K}_A(\alpha, \lambda(\alpha_K, n), n)$ distributed data with $\lambda(\alpha_K, n)$ as previously defined. From this sample α_G, the homogeneity parameter of the $\mathcal{G}_A^0(\alpha_G, \gamma(\alpha_G, n), n)$ distribution is estimated using the maximum likelihood estimator presented in eq. (5). If the random variable Z is $\mathcal{G}_A^0$ distributed then its cumulative distribution function can be evaluated using eq. (3).

5 Monte Carlo experience

A Monte Carlo experience was carried out generating $\mathcal{K}_A$ distributed data which were fitted with a $\mathcal{G}_A^0$ distribution. The values of the parameters α_K and α_G were estimated by the maximum likelihood method and by the $\frac{1}{2}$order moment estimator method. The goodness of fit was evaluated using the p-value of the χ^2 test.

For a number R of replications, the following steps were performed:

1. For each $\alpha_K \in [4, 12]$ $\mathcal{K}_A$ distributed samples were generated.

2. For each of these $\mathcal{K}_A$ distributed samples the roughness parameter α_G and the parameter γ of the $\mathcal{G}_A^0$

distribution were estimated.

3. The goodness of fit was evaluated using the p-value of the test.

The $\mathcal{K}_A$ distributed samples were generated for values of $\alpha_K \in [4, 12]$ with $R = 100$, 1000 and 10000 replications, sample sizes $T = 1000$ and 10000 and significance level 0.01. As a particular case, the figures for $\alpha_K = 4$ are shown in Table 2: for each number of replications (R) and each sample size (T) the mean of the estimated parameters $\widehat{\alpha}_G$ is presented, along with the mean square error of the estimation (mse) and with the mean significance level of the cross-fit ($r_{0.01}$). This shows that there is no reason to suppose that the $\mathcal{K}_A$ and $\mathcal{G}_A^0$ distributions are different at the proposed significance level. Table 3 shows the values corresponding to the mean value of the estimated α_G with $\alpha_K \in [5, 10]$ and $n = 1$.

R	T	α_G	mse	$r_{0.01}$
100	10000	−4.50	0.13	11.0
1000	1000	−4.78	1.80	1.5
1000	10000	−4.53	0.14	10.0
10000	1000	−4.79	2.08	1.3
10000	10000	−4.52	0.13	12.0

Table 2: Correspondence between $\alpha_K = 4$ and $n = 1$.

α_K	α_G	mse	$r_{0.01}$
5	−5.53	0.25	3.4
6	−6.55	0.50	2.2
7	−7.61	0.95	1.0
8	−8.73	1.93	1.1
9	−9.69	2.44	1.3
10	−10.83	4.61	0.7
11	−11.98	6.80	0.7

Table 3: Correspondence for $R = 1000$, $T = 10000$ and $n = 1$.

6 Comparison of classifications

Image processing and analysis requires the use of filtering, segmentation and classification operations. For this reason, the influence the change of model exerts on the classification results will be analysed in this section.

An image with regions of different degrees of homogeneity will be classified. Even though data are $\mathcal{K}_A$ distributed, the image will be classified with both the $\mathcal{G}_A^0$ and the $\mathcal{K}_A$ model, and the results will be compared. A real SAR image containing areas of several degrees of homogeneity will be used for the proposed models.

A 400×400 subimage, taken from a 1600×2400 JERS-1, L band, HH polarisation image with an estimated number of looks equal to 2.95, corresponding to the Floresta Nacional do Tapajós (State of Pará, Brazil, shown in Fig. 1), was classified using the , the $\mathcal{K}_A$ and the $\mathcal{G}_A^0$ models. It was observed that using only the $\mathcal{G}_A^0$ distribution, the classification results are statistically as accurate as the ones obtained with the two other models. This occurs in homogeneous as well as in heterogeneous areas.

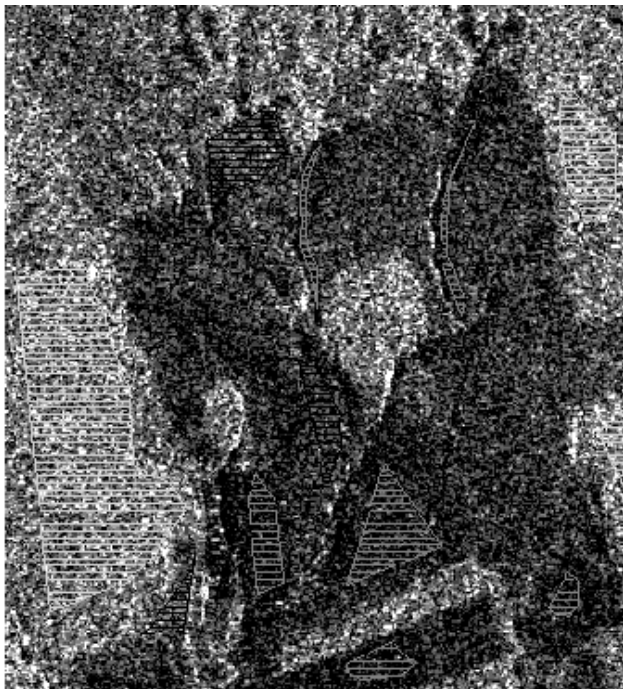

Figure 1: JERS-1 image with training regions corresponding to clear-cut (dark), cut with sec-ondary growth (intermediate), and forest (light).

These classifications were obtained using Maximum Likelihood followed by the ICM (Iterated Conditional Modes) algorithm. Maximum Likelihood is a supervised classification method widely used in remote sensing. The ICM method [1], is an iterative, deterministic local optimisation procedure.

Three classes were used: clear-cut (CC, dark grey), cut with secondary growth (CSG: medium grey) and forest (F: light grey). The training regions are shown in Fig. 1.

The estimated parameter values for the three classes under the $\mathcal{K}_A$ were: CC: $\widehat{\alpha}_K = 16.6$, $\widehat{\lambda} = 0.0041$,

CSG: $\widehat{\alpha}_K = 13.76, \widehat{\lambda} = 0.0015$ and F: $\widehat{\alpha}_K = 12.56, \widehat{\lambda} = 0.00083$. The estimated parameter values for the $\mathcal{G}_A^0$ model were: CC: $\widehat{\alpha}_G = -19.42$, $\widehat{\gamma} = 74706.26$, CSG: $\widehat{\alpha}_G = -14.17$, $\widehat{\gamma} = 119671.72$ and F: $\widehat{\alpha}_G = -12.97$, $\widehat{\gamma} = 178364.07$. It can be observed CC and CSG correspond to relatively homogeneous areas, while F is more heterogeneous than the former two. Confusion matrixes were calculated in order to compare the results.

In the first case, data of the training areas corresponding to the three proposed classes were modelled with the $\Gamma^{1/2}$ distribution. Figure 2 shows the classification result of applying ML; then ICM was applied. Since both ML and ICM results are visually identical under the three considered models, only ML under the $\Gamma^{1/2}$ model (Fig. 2) and ICM under the $\mathcal{G}_A^0$ model (Fig. 3) are shown. Since ICM classification always represents an improvement over ML, we chose it to perform the comparison among the three models.

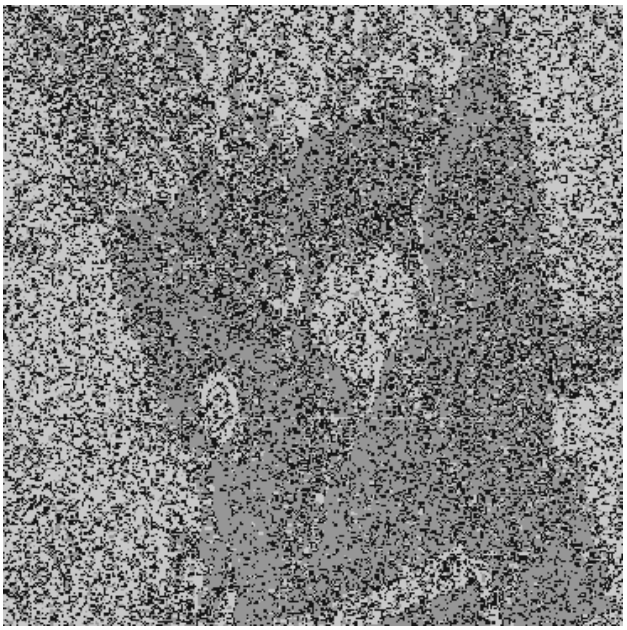

Figure 2: JERS-1 classification by Maximum Likelihood, under the $\Gamma^{1/2}$ distribution model.

Table 4 shows the confusion matrix for the ICM classification under the model. Here, the percentage of pixels classified in each class, the overall accuracy of the classification and the corresponding κ coefficient of agreement are reported.

Table 5 and Table 6 show the confusion matrixes of the ICM classifications under the $\mathcal{K}_A$ and $\mathcal{G}_A^0$ models, and they lead to the conclusion that the $\mathcal{G}_A^0$ distribution does not deteriorate the classifications.

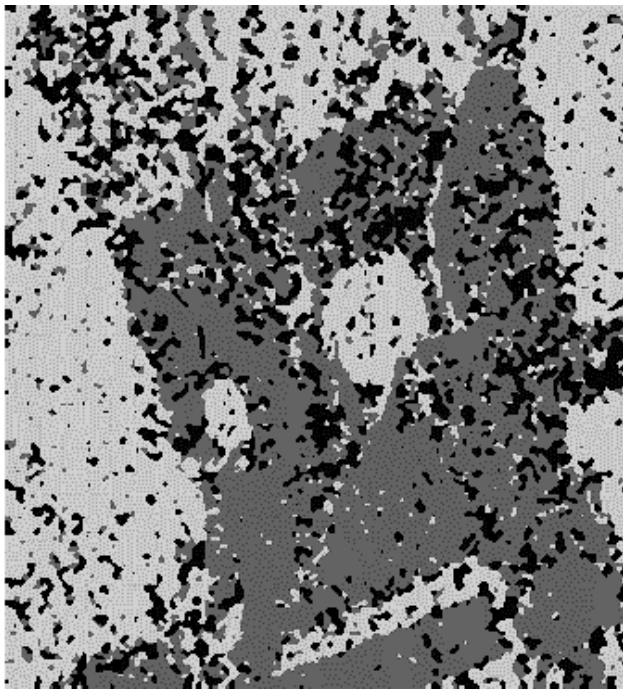

Figure 3: JERS-1 classification using the ICM method, under the $\mathcal{G}_A^0$ distribution model.

Accuracy 85.8%, $\kappa = 0.81$				
CC	39.054	1.74	2.755	43.549
CSG	0.187	40.317	1.692	42.195
F	1.007	6.820	6.430	14.256
Total	40.247	48.876	10.876	100.000

Table 4: Confusion matrix (percent) of the $\Gamma^{1/2}$ model.

7 The $\mathcal{G}_A^0$ as the true model

As the ground truth can be characterised by the parameters α and γ, their estimation for each pixel may lead to estimated parameters maps that, in turn, can be used as the input for classification methods, among other applications. These features convey important information that helps the understanding of the scene.

Several parameter estimation techniques are available, being the most remarkable ones those based on maximum likelihood, on sample moments or substitution method (MO for short), on order statistics and on data transformations (see [5]). In this work MO estimation is used. To estimate α and γ it is necessary to estimate two moments. In this work moments of order 1/2 and 1, denoted as $m_{1/2}$ and m_1 respectively, will

Accuracy 88.0%, $\kappa = 0.84$				
CC	38.907	1.232	2.460	42.599
CSG	0.234	45.25	2.408	45.761
F	1.006	6.820	6.009	11.640
Total	40.247	48.876	10.876	100.000

Table 5: Confusion matrix (percent) of the $\mathcal{K}_A$ model.

Accuracy 85.1%, $\kappa = 0.81$				
CC	39.054	1.740	2.755	43.549
CSG	0.187	40.317	1.692	42.195
F	1.007	6.820	6.430	14.256
Total	40.247	48.876	10.876	100.000

Table 6: Confusion matrix (percent) of the $\mathcal{G}_A^0$ model.

be used. These moments are given by

$$m_{1/2} = \left(\frac{\gamma}{n}\right)^{1/4} \frac{\Gamma(-\alpha - 1/4)\Gamma(n + 1/4)}{\Gamma(-\alpha)\Gamma(n)}, \alpha < -1/4$$

$$m_1 = \left(\frac{\gamma}{n}\right)^{1/2} \frac{\Gamma(-\alpha - 1/2)\Gamma(n + 1/2)}{\Gamma(-\alpha)\Gamma(n)}, \alpha < -1/2$$

Using these relations $\hat{\alpha}$ can be determined as the solution of the equation

$$g(\hat{\alpha}) - \zeta = 0, \tag{6}$$

where

$$g(\hat{\alpha}) = \frac{\Gamma^2(-\hat{\alpha} - 1/4)}{\Gamma(-\hat{\alpha})\Gamma(-\hat{\alpha} - 1/2)}$$

and

$$\zeta = \frac{\hat{m}_{1/2}^2 \Gamma(n)\Gamma(n + 1/2)}{\hat{m}_1 \Gamma(n + 1/4)},$$

and then replacing the value of $\hat{\alpha}$ in one of the expressions of the moments to find $\hat{\gamma}$.

It can be seen that, since $g(\hat{\alpha})$ converges asymptotically to 1 and since as ζ is a random variable that can take values greater than one, there are cases for which eq. (6) does not have a solution. The lower the value of $\hat{\alpha}$, the higher the chances of not being a solution for eq. (6). For our application, a technique to overcome this problem is presented.

One of the most important available manners to assess the adequacy of a theory to reality is the use of stochastic simulation. In the following, an experiment involving simulation and estimation will be shown, aiming at illustrating the model and tools presented. This simulated image will also be used to test the performance of the forthcoming classification procedure.

It is immediate to obtain outcomes from $\mathcal{G}_A^0$ distributed random variables, since algorithms for sampling from Γ distributions are available. In order to obtain outcomes of the random variable $\mathcal{G}_A^0(\alpha, \gamma, n)$ it is enough to return the outcome of $\sqrt{U/V}$, where $U \sim \Gamma(n, n)$ is independent of $V \sim \Gamma(-\alpha, \gamma)$.

An idealised scene with four regions will be used here. The regions represent heterogeneous ($\alpha = -5$) and extremely heterogeneous ($\alpha = -1.5$) areas. Two values of γ, namely $\gamma \in \{2.10^5, 4.10^5\}$, were used for each α. The result of this simulation is shown in Fig. 4. The $n = 1$ case is chosen for this study since it corresponds to the noisiest (and hardest to classify) class of amplitude images: single look data.

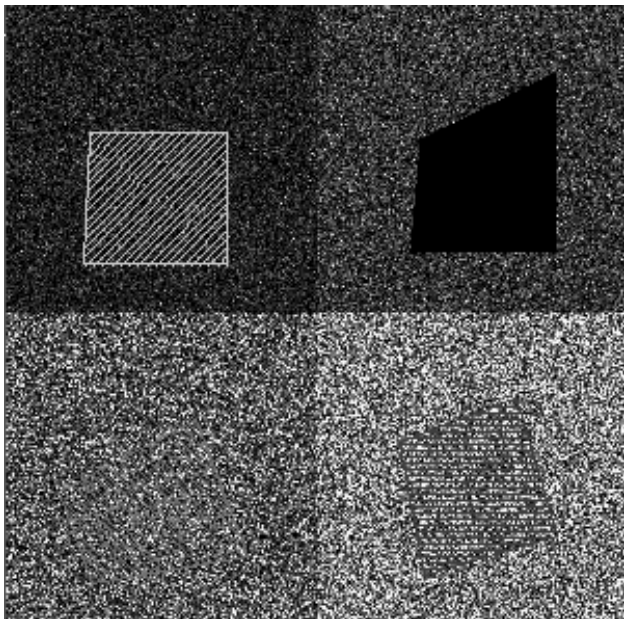

Figure 4: Synthetic image with four areas (top to bottom and left to right): ($\alpha = -5, \gamma = 2.105$), ($\alpha = -5, \gamma = 4.105$), ($\alpha = -1.5, \gamma = 2.105$) and ($\alpha = -1.5, \gamma = 4.105$).

The image presented in Fig. 4 was then analysed within the context of Gaussian maximum likelihood, i.e., representative samples from each region were taken in order to estimate the mean and variance of normal distributions. The results of this training step were the estimated parameters $\hat{\mu} = 193.10$ and $\hat{\sigma} = 115.42$ for the upper left class (C1 originally $\alpha = -5$, $\gamma = 2.105$), $\hat{\mu} = 272.16$ and $\hat{\sigma} = 163.15$ for the upper right class (C2 originally $\alpha = -5$, $\gamma = 4.105$), $\hat{\mu} = 444.91$ and $\hat{\sigma} = 405.11$ for the lower left class (C3 originally $\alpha = -1.5$, $\gamma = 2.105$), and $\hat{\mu} = 634.72$ and $\hat{\sigma} = 606.84$ for the lower right class (C4 originally $\alpha = -1.5$, $\gamma = 4.105$). The densities of these distributions are used to determine the classification rule that, in turn, yielded the

map presented in Fig. 5. This map is clearly dominated by the upper left class, and it is unacceptable for any practical application.

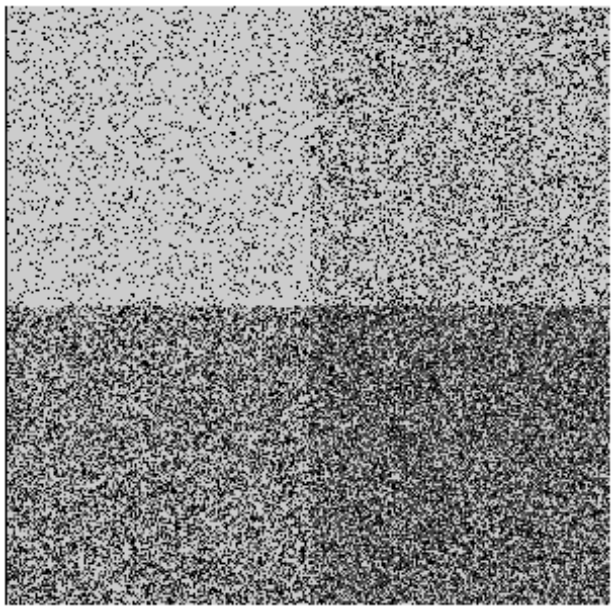

Figure 5: Gaussian maximum likelihood classification of the simulated image (Fig. 4).

Fig. 6 shows the result of applying the estimators for α and γ to the image shown in Fig. 4. The image on the left was obtained by locally estimating α and the image on the right is the result of locally estimating γ. In both cases, a 7×7 sliding window was used.

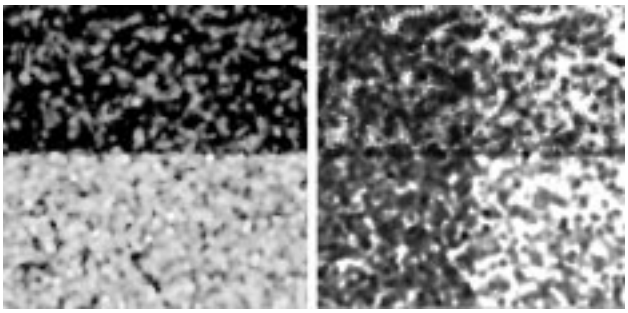

Figure 6: Estimators of the roughness (α, left) and scale (γ, right) on 7×7 windows over the simulated image (Fig. 4).

As previously stated, there are cases where no estimated value is available (i.e. where eq. (6) has no solution). The co-ordinates of such pixels are marked, and their corresponding estimated values are derived calculating the median of the observed estimations in a 11×11 window around it. This rarely happens in extremely heterogeneous areas, and the number of pixels where this phenomenon is observed depends on the number of observations used in the estimation (the bigger the window size, the less pixels without estimation).

From Fig. 6 left, it is possible to conclude that the estimator $\hat{\alpha}$, based on the m_1 and $m_{1/2}$ moments, distinguishes between heterogeneous and extremely heterogeneous areas, since the upper half of the image (where original observations were generated with $\alpha = -1.5$) is darker than the lower half (where $\alpha = -5$ was used in the simulation). It is also noticeable that $\hat{\alpha}$ does not discriminate between right and left halves of Fig. 4, where the roughness parameter a was kept constant.

Fig. 6 right leads to the conclusion that the estimator $\hat{\gamma}$ is able to separate **both** regions where the scale changes (right and left halves) and areas where different roughness are observed (upper and lower halves). This last discrimination is performed to a lesser extent than the estimator $\hat{\alpha}$ does (notice that the contrast of the right image is smaller than the left one). Then the parameter γ can be successfully used as a tool for the discrimination among different types of targets, particularly in those regions where the parameter α cannot perform it. This can be very useful in cases of areas with the same roughness but different brightness.

The joint behaviour of $\hat{\alpha}$ and $\hat{\gamma}$ suggests that they may form a good bi-level feature for the classification of SAR data. The image shown in Fig. 4 will be classified, using the features shown in Fig. 6.

Fig. 7 shows the result of applying Gaussian maximum likelihood classification to the bi-level image shown in Fig. 6. The confusion matrix of this classification is shown in Table 7, where the training samples were used as ground truth. This confusion matrix allows us to say that the classification is excellent.

Result\True	C_1	C_2	C_3	C_4
C_1	96.81	1.22	3.57	0.00
C_2	0.00	93.28	0.00	0.43
C_3	2.15	0.00	96.11	0.00
C_4	1.04	5.50	0.32	99.57

Table 7: Confusion matrix from the classification of the simulated four classes image.

Fig. 8 shows an E-SAR image, obtained in HH polarisation and L-band, slant range and calibrated without elevation slope correction. This is an airborne

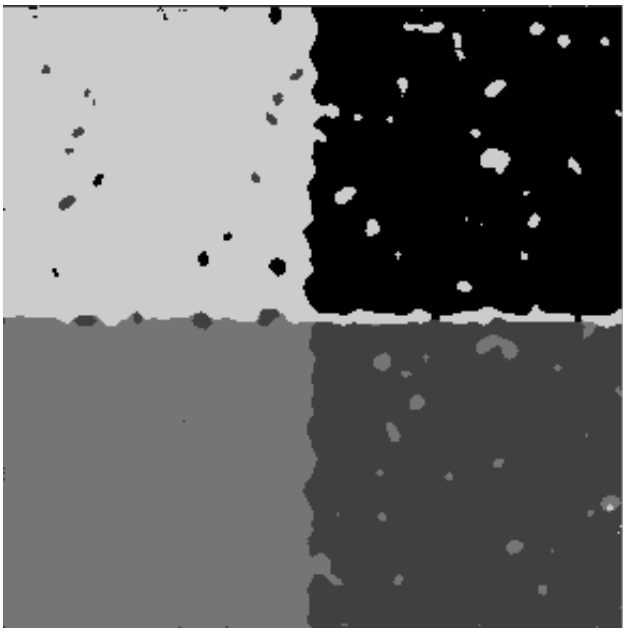

Figure 7: Gaussian maximum likelihood classification of the estimated parameters images.

sensor belonging to the German Aerospace Institute (DLR, Oberpfaffenhofen). The area shown here shows a homogeneous region (pasture), surrounded by extremely heterogeneous return corresponding to an urban area. A Gaussian maximum likelihood classification was tried in this image using the two aforementioned classes. The result was the complete dominion of the urban class over the other, i.e., almost all the pixels were classified as belonging to the urban type.

Fig. 9 shows the extracted features from the original image, the roughness parameters to the left and the scale parameters to the right. It is noticeable that the homogeneous areas appear as dark spots in the former, and a as light region in the latter. This discrimination is inverted for the scale image. This pair of images was then used as a bi-level feature image for a Gaussian maximum likelihood classification, and the result is presented in Fig. 10. The confusion matrix (in percentages) of this classification is presented in Table 8, where the training samples were used as ground truth. This confusion matrix allow us to say that this classification is excellent.

The interpretability of the parameters, as already commented, allow the verification of the labels used for the classes observed in Fig. 8. The data from these classes that were used as training samples produced $\hat{\alpha} = -12.5$ in the area labelled as "pasture" and

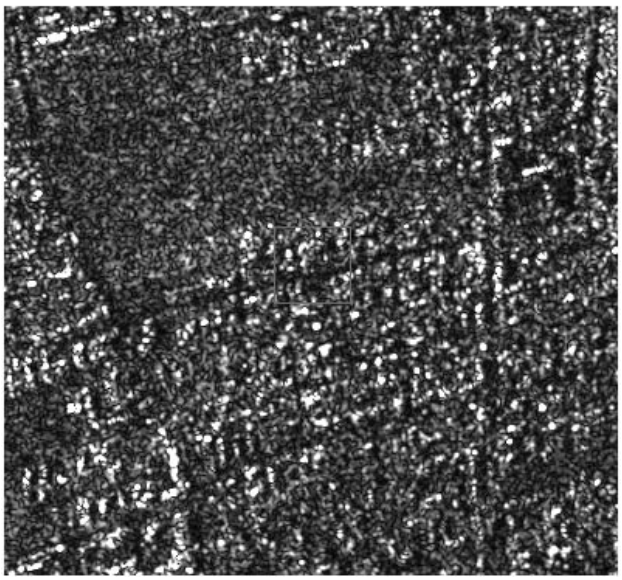

Figure 8: E-SAR image over Gilching.

$\hat{\alpha} = -1.6$ for the one called "urban". Pasture areas are often seen as homogeneous whilst the return from urban ones is extremely heterogeneous.

The other image analysed by means of the extraction of parameter maps and Gaussian maximum likelihood is the one already shown in Fig. 1. Three regions are distinguishable: clear cut (seen as homogeneous data, dark grey), cut with secondary growth (less homogenous data, intermediate grey) and virgin forest (heterogeneous data, light grey).

Again, a Gaussian maximum likelihood classification of the raw data yielded a completely unacceptable result, so parameter maps were extracted. These maps are shown in Fig. 11, the roughness to the left and the

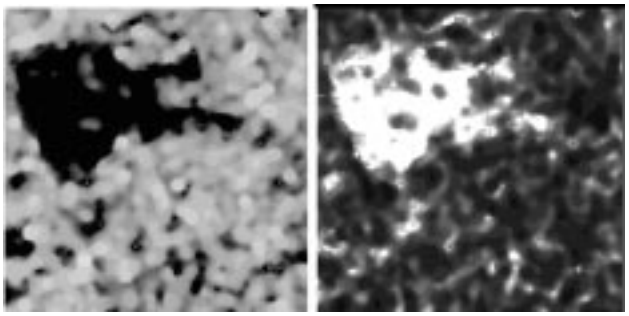

Figure 9: Estimated roughness (left) and scale (right) parameters data obtained from the E-SAR image.

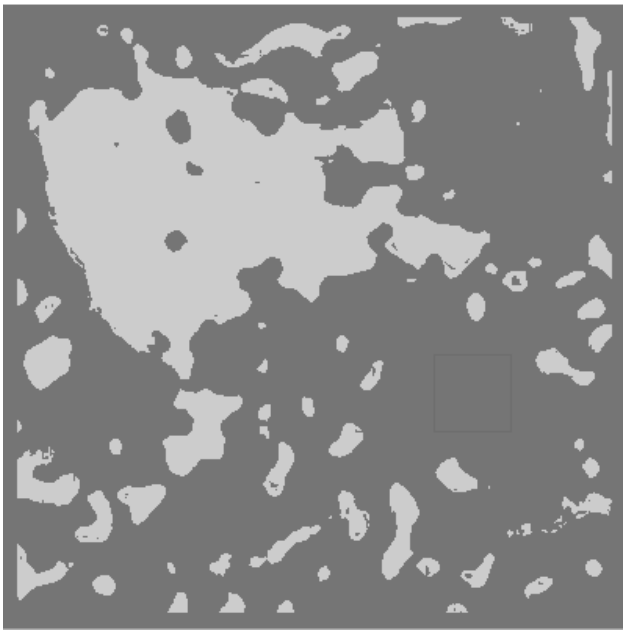

Figure 10: Gaussian maximum likelihood classification of the estimated parameters images.

Result\True	Pasture	Urban
Pasture	97.62	1.84
Urban	2.38	98.16

Table 8: Confusion matrix from the classification of the Gilching image.

scale to the right. The estimated parameters in the training areas were ($\hat{\alpha} = -19.42$, $\hat{\gamma} = 74706.26$) in the clear cut, ($\hat{\alpha} = -14.17$, $\hat{\gamma} = 119671.72$) in the secondary growth and ($\hat{\alpha} = -12.97$, $\hat{\gamma} = 178364.07$) in the virgin forest. It can be observed, thus, that the two first classes (clear cut and cut with secondary growth) correspond to relatively homogeneous areas, while the virgin forest is more heterogeneous than the other two.

The Gaussian maximum likelihood classification of the data presented in Fig. 11 is shown in Fig. 12, and the confusion matrix (in percentage) of this classification, using training data as the truth, is shown in Table 9. Again, the classification is quantitatively excellent.

8 Conclusions

In this work, the possibility of substituting the $\mathcal{K}_A$ distribution by the $\mathcal{G}_A^0$ distribution has been presented. The importance of this substitution resides in the theoretical and computational tractability of the second

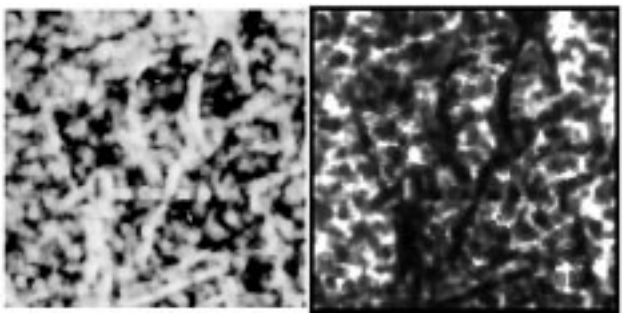

Figure 11: Gaussian maximum likelihood classification of the estimated parameters images.

Result\True	Clear Cut	Sec. Growth	Forest
Clear Cut	94.44	6.56	0.01
Sec. Growth	5.56	87.44	3.53
Forest	0.00	6.00	96.46

Table 9: Confusion matrix from the classification of JERS image.

one over the first one.

It has been shown in a Monte Carlo experience that, for a given value of the roughness parameter α_K, the replications of the estimated values of α_G are close to the theoretical value found through the minimisation of the distance between distributions. Thus, the hypothesis of $\mathcal{K}_A$ distributed data can be substituted by the hypothesis of $\mathcal{G}_A^0$ distributed data. This was demonstrated for the roughness parameter α_K varying within the interval $[4, 12]$.

In the first practical application presented here the $\Gamma^{1/2}$, the $\mathcal{K}_A$ and the $\mathcal{G}_A^0$ have been considered as possible models for SAR data. Maximum likelihood classifications followed by the ICM algorithm were carried out for the three models in a JERS-1 image, containing areas of several degrees of homogeneity: clear-cut (CC), cut with secondary growth (CSG) and forest (F). The obtained results show that the classification performance does not deteriorate when the $\mathcal{G}_A^0$ distribution is adopted as the true model.

In a second practical application, estimation of the parameters of the $\mathcal{G}_A^0$ distribution based on moments was used to derive features maps that were used as the input of Gaussian maximum likelihood classification. This was done on a simulated 1-look SAR image, on a 1-look, HH polarisation, L-band E-SAR image and on the same JERS-1 image used in the first practical application. The results, both using simulated and real SAR images are excellent. The computational effort to

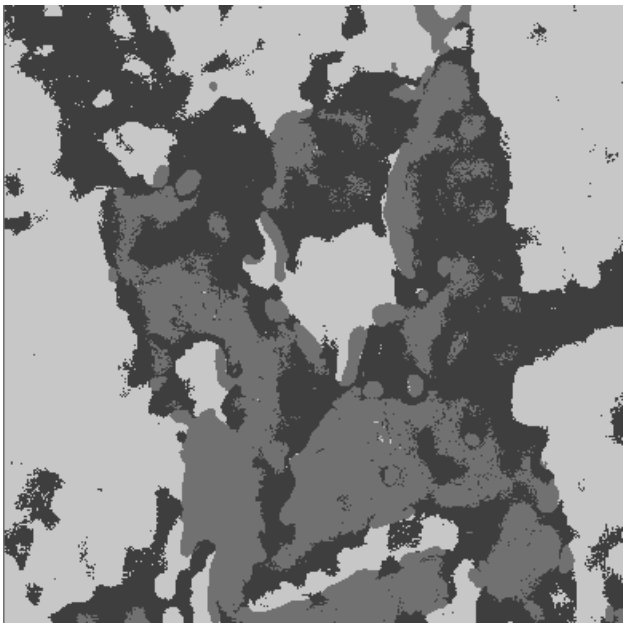

Figure 12: Gaussian maximum likelihood classification of estimated maps.

derive the parameter maps is affordable.

We can then say that the $\mathcal{G}_A^0$ distribution is quite a good model for SAR data, that its parameters have relevant and immediate interpretation and that the estimation of these parameters, which is a simple computational task, allows the derivation of features that can be used to obtain excellent classifications with Gaussian maximum likelihood.

Given these results and the theoretical as well as the practical advantages of the $\mathcal{G}_A^0$ distribution over the $\mathcal{K}_A$ distribution, we conclude that the former can be adopted as the true model for SAR data.

Acknowledgments

This work was partially supported by grants from CNPq (Proc. 523469/96-9) and FACEPE (APQ 0707-1.03/97).

References

[1] J. Besag, "On the statistical analysis of dirty pictures (with discussion). In: *Journal of the Royal Statistical Society B*, 48(3)259–302, (1986).

[2] A. C. Frery, H.-J. Müller, C. C. F Yanasse, S. J. S. Sant'Anna, "A model for Extremely Heterogeneous Clutter. In: *IEEE Transactions on Geoscience and Remote Sensing*, 35(3)648–659, (1997).

[3] S. D. Gordon, J. A. Ritcey, "Calculating the K distribution by saddlepoint integration". In: *IEE Proceedings in Radar, Sonar and Navigation*, 142(4)162–165, (1995).

[4] I. R. Joughin, D. B. Percival, D. P. Winebrenner, "Maximum likelihood estimation of K distribution parameters for SAR data. In: *IEEE Transactions on Geoscience and Remote Sensing*, 31(5)989–999, (1993).

[5] M. Mejail, *The GA0 distribution in the modelling and analysis of SAR images*. PhD Thesis in Computer Science, Universidad de Buenos Aires, Argentina, 1999.

[6] C. C. F Yanasse, A. C. Frery, S. J. S. Sant'Anna, "Stochastic distributions and the multiplicative model: relations, properties, estimators and applications to SAR image analysis. Technical Report 5630-NTC/318, INPE, So Jos dos Campos, SP, Brazil, (1995).

[7] C. C. F Yanasse, A. C. Frery, S. J. S. Sant'Anna, P. F. Hernandez, L. V. Dutra, "Statistical analysis of SAREX data over Tapajós, Brazil. In: *SAREX-92: South American Radar Experiment*, M. Wooding and E. Attema, eds. pp. :25–40, ESA, Paris, (1993).