

INPE-5645-PRE/1826

STATISTICAL CHARACTERIZATION OF SAR DATA: THE MULTIPLICATIVE MODEL  
AND EXTENSIONS

Alejandro C. Frery  
Corina da Costa Freitas Yanasse  
Sidnei J. S. Santa'Anna

Simposio Latinoamericano de Especialistas en Percepción Remota, Puerto  
Vallarta, México, 1995

INPE  
São José dos Campos  
1995

# STATISTICAL CHARACTERIZATION OF SAR DATA: THE MULTIPLICATIVE MODEL AND EXTENSIONS

Alejandro C. Frery  
Corina da Costa Freitas Yanasse  
Sidnei J. S. Sant'Anna  
INPE — Instituto Nacional de Pesquisas Espaciais  
DPI — Divisão de Processamento de Imagens  
Avenida dos Astronautas, 1758  
12227-010 São José dos Campos  
SP - Brazil  
{frery;corina;sidnei}@dpi.inpe.br

## ABSTRACT

In this paper a comprehensive account of the distributions associated to three SAR image formats (namely, complex, intensity and amplitude) and types of classes (homogeneous, heterogeneous and extremely heterogeneous) is presented, along with their main properties and relations. The backscatter of extremely heterogeneous targets is modelled by the square root of generalized inverse Gaussian distribution, that yields to an extension of K distributions for the return. Modelling with K distributions is applied to SAR data over amazonian areas, showing the discriminatory capability of the methodology. For the sake of completeness, some other distribution that do not fit in the multiplicative model (but that have proved their usefulness) are discussed.

## 1 INTRODUCTION

The precise knowledge of the statistical properties of SAR data plays a central role in image processing and understanding. These properties can be used to discriminate types of land use and to develop specialized filters for speckle noise reduction, among other applications. Several studies have been conducted in order to relate physical features and statistical properties of SAR data and, in order to do this, some hypothesis and distributions are considered.

The multiplicative model has been widely used in the modelling, processing and analysis of synthetic aperture radar images. The multiplicative hypothesis states that the return results from the product between the speckle noise and the terrain backscatter. The distributions associated to the speckle noise can be assumed known since, when the number of elementary backscatters is large enough and no one dominates the return, among other factors, they either are the bivariate normal or they only depend on the number of looks. Several distributions could be used for the backscatter, aiming at the modelling of different types of classes and their characteristic degrees of homogeneity; for instance, for some sensor characteristics (wavelength, angle of incidence, polarization, etc.), pasture is more homogeneous than forest which, in turn, is more homogeneous than urban areas. Most

distributions for the (amplitude) backscatter do not yield to closed-form distributions for the return, being a constant (homogeneous), a square root of Gamma (heterogeneous) and a square root of generalized inverse Gaussian distribution (extremely heterogeneous areas, resp.) important exceptions.

Common distributional hypothesis for 1-look return data and homogeneous targets are the Exponential and Rayleigh distributions, for quadratic and linear detections respectively. When the observed region cannot be assumed as homogeneous, other distributions are considered. Among these, the  $K$  distributions have received a great deal of attention in the literature (Jakeman, 1973; Jakeman and Pusey, 1976; Jao, 1984).

This paper reviews the derivation of those distributions arising in the multiplicative model under several hypothesis, and extends classical results with the proposal of distributions that have the  $K$  distributions as particular cases. These distributions arise using the square root of generalized inverse Gaussian distribution as the model for the amplitude backscatter, a distribution here proposed to model extremely heterogeneous targets.

This work is organized in the following manner: all the cases (complex, intensity and amplitude) for the speckle noise are treated separately; distributions that model the terrain backscatter (amplitude and intensity cases) are then reviewed, and the generalized inverse Gaussian distribution is introduced. Once the speckle noise and the backscatter are characterized, attention is devoted to the distributions induced for the return. Other distributions that do not arise from the multiplicative model, but that have proved their usefulness in SAR data modelling and analysis, are also presented. Finally, applications to SAR data analysis are presented.

## 2 THE MULTIPLICATIVE MODEL AND THE SPECKLE NOISE

Speckle noise is always associated to coherent-illuminated scenes, such as those obtained by microwaves, laser, ultrasonography, etc. This kind of noise appears due to interference phenomena between the incident and reflected signals. This kind of noise makes visual and automatic interpretation a difficult task, though it may carry important information. As will be seen in this section, images suffering from speckle noise should not be treated with the usual additive-noise derived tools (Wiener filter, for instance), since speckle corrupts the signal in a multiplicative manner and in the amplitude and intensity formats it is non-Gaussian.

The multiplicative model is a common framework used to explain the statistical behavior of data obtained with coherent illumination. It assumes that the observations within this kind of images are the outcome of the product of two independent random variables: one ( $X$ ) modelling the terrain backscatter, and other ( $Y$ ) modelling the speckle noise. The former is always considered real and positive, while the latter could be complex (if the considered image is in complex format) or positive real (intensity and amplitude formats).

Therefore, the observed value is the outcome of the random variable defined by the product  $Z = X \cdot Y$ . In order to make a clear distinction between the aforementioned formats, the indexes “ $C$ ”, “ $I$ ” and “ $A$ ” will be used for the complex, intensity and amplitude cases, respectively. Vectors will be explicitly denoted in boldface.

Complex speckle is usually assumed to have a bivariate normal distribution, with independent identically distributed components having zero mean and variance  $1/2$ . These marginal distributions are denoted here as  $\mathcal{N}(0, 1/2)$ ; therefore,  $\mathbf{Y}_C = (Y_{\Re}, Y_{\Im}) \sim \mathcal{N}^2(\mathbf{0}, \mathbf{1}/2)$  denotes the distribution of the pair. This assumption is justified considering that every resolution cell is formed by a large number of point scatters and, therefore, that the returned signal might be regarded as a random walk. The resolution cell in this case can be modelled

as a collection of  $N$  point scatters, and the received signal as a sum of randomly phased phasors. If the  $k$ -th elementary phasor is expressed as  $v_k e^{i\phi_k}$ , where  $v_k$  is its amplitude and  $\phi_k$  its instantaneous phase, then the received signal is

$$\sum_{1 \leq k \leq N} v_k \exp(i\phi_k) = Y_A \exp(i\phi),$$

where  $Y_A$  is the resulting amplitude and  $\phi$  is the (instantaneous) resulting phase.

Decomposing the received signal into real and imaginary parts  $Y_{\Re}$  and  $Y_{\Im}$ , where

$$\begin{aligned} Y_{\Re} &= Y_A \cos \phi = \sum_{1 \leq k \leq N} v_k \cos(\phi_k) \\ Y_{\Im} &= Y_A \sin \phi = \sum_{1 \leq k \leq N} v_k \sin(\phi_k), \end{aligned}$$

it is possible to derive their distributional properties making the following assumptions for every  $k$ :

1. the amplitude  $v_k$  and the phase  $\phi_k$  of the  $k$ -th elementary phasor are independent of each other and of the amplitudes and phases of all other elementary phasors.
2.  $\phi_k$  is uniformly distributed over the range  $[0, 2\pi)$ .
3.  $N$  is large and the Lindeberg conditions hold.

In this case it can be proved that  $Y_{\Re}$  and  $Y_{\Im}$  are independent and identically distributed as  $\mathcal{N}(0, \sigma^2)$ , with  $\sigma^2 = \frac{1}{2} \sum_{1 \leq k \leq N} E(v_k^2)$ . A convenient parametrization consists of setting  $\sigma^2 = 1/2$ , since pure speckle is being modelled without regarding terrain properties and it is desirable to have an intensity speckle mean equal to 1.

The aforementioned signal amplitude ( $Y_A$ ) and phase ( $\phi$ ) are related to the real and imaginary components ( $Y_{\Re}$  and  $Y_{\Im}$ ) through  $Y_A = \sqrt{Y_{\Re}^2 + Y_{\Im}^2} = \|\mathbf{Y}_C\|$  and  $\phi = \arctan(Y_{\Im}/Y_{\Re})$ . It is usually employed the term *intensity* for  $Y_A^2 = Y_I$ . It is easy to prove that under the above conditions  $Y_I$  has an Exponential distribution,  $Y_A$  has a Rayleigh distribution and  $\phi$  is uniformly distributed over the range  $[0, 2\pi)$ .

Multilook intensity speckle appears by taking the average over  $n$  independent samples of  $Y_I = \|\mathbf{Y}_C\|^2$  leading, thus, to a Gamma distribution denoted here as  $Y_I \sim \Gamma(n, n)$  and characterized by the density

$$f_{Y_I}(y) = \frac{n^n}{\Gamma(n)} y^{n-1} \exp(-ny), \quad y, n > 0.$$

Multilook amplitude speckle results from taking the square root of the multilook intensity speckle and, therefore, has a square root of Gamma distribution, denoted here as  $Y_A \sim \sqrt{\Gamma}(n, n)$  and characterized by the density

$$f_{Y_A}(y) = \frac{2n^n}{\Gamma(n)} y^{2n-1} \exp(-ny^2), \quad y, n > 0.$$

Though the number of looks  $n$  should, in principle, be an integer, seldom this is the case when this quantity is estimated from real data due to, among other reasons, the fact that the mean is taken over correlated observations. It is therefore interesting to call  $n$  the *equivalent* number of looks, and estimation procedures to determine this quantity are discussed in Yanasse et al. (1994).

### 3 AMPLITUDE BACKSCATTER

Three main models for the amplitude backscatter yield to closed-form distributions for the return: a constant (used to model homogeneous areas), a square root of Gamma distribution (used to model heterogeneous areas) and the square root of a generalized inverse Gaussian distribution (proposed here as a model for extremely heterogeneous areas). The introduction of the third distribution was originated in the analysis presented in Yanasse et al. (1994, 1995), since there it was noticed that the observations from some areas were heterogeneous to an extent that even  $K$  distributions could not take account of. The definition of a constant as a random variable is used to fit every case into a completely stochastic model.

The backscatter from a completely homogeneous area is said to have the distribution of a constant, here denoted as  $X_A \sim \mathcal{C}(\sqrt{\beta})$ , if its cumulative distribution function is given by  $F_{X_A}(x) = 1_{[\beta^{1/2}, \infty)}(x)$ , where  $\beta > 0$  and  $1_A$  denotes the indicator function of the set  $A$ .

The backscatter from a heterogeneous area is said to have a square root of Gamma distribution with parameters  $\alpha, \lambda > 0$ , here denoted as  $X_A \sim \sqrt{\Gamma}(\alpha, \lambda)$ , if its density is given by

$$f_{X_A}(x) = \frac{2\lambda^\alpha}{\Gamma(\alpha)} x^{2\alpha-1} \exp(-\lambda x^2), \quad x, n > 0.$$

This hypothesis is mainly based on empirical evidence and on the fact that the product of two square root of Gamma distributed random variables has an explicit density, though some theoretical results relating backscatter with birth-and-dead processes can be found in Caves (1993). Its  $r$ -th order moments are given by

$$E(X_A^r) = \frac{\Gamma(\alpha + r/2)}{\lambda^{r/2} \Gamma(\alpha)}.$$

The backscatter from extremely heterogeneous areas is said to have a square root of generalized inverse Gaussian distribution, here denoted as  $X_A \sim \mathcal{N}^{-1/2}(\alpha, \gamma, \lambda)$ , if its density is given by

$$f_{X_A}(x) = \frac{(\lambda/\gamma)^{\alpha/2}}{K_\alpha(2\sqrt{\lambda\gamma})} x^{2\alpha-1} \exp\left(-\frac{\gamma}{x^2} - \lambda x^2\right), \quad x > 0,$$

where  $K_\alpha$  denotes the modified Bessel function of the third kind and order  $\alpha$ , and the parameters space is given by

$$\begin{cases} \gamma > 0, & \lambda \geq 0 & \text{if } \alpha < 0 \\ \gamma > 0, & \lambda > 0 & \text{if } \alpha = 0 \\ \gamma \geq 0, & \lambda > 0 & \text{if } \alpha > 0 \end{cases} \quad (1)$$

Its  $r$ -th order moments are given by

$$E(X_A^r) = \left(\frac{\gamma}{\lambda}\right)^{r/4} \frac{K_{\alpha+r/2}(2\sqrt{\gamma\lambda})}{K_\alpha(2\sqrt{\gamma\lambda})}$$

Important limiting properties that relate the aforementioned distributions can be posed in the following manner

$$\mathcal{N}^{-1/2}(\alpha, \gamma, \lambda) \xrightarrow[\gamma \downarrow 0]{\mathcal{D}} \sqrt{\Gamma}(\alpha, \lambda) \xrightarrow[\alpha/\lambda \rightarrow \beta]{\mathcal{D}} \mathcal{C}(\sqrt{\beta}), \quad (2)$$

where “ $\mathcal{D}$ ” denotes the convergence in distribution of the associated random variable. The properties stated in relations (2) assure the unicity of the model, since every type of amplitude backscatter can be treated as the outcome of a special case of  $X_A \sim \mathcal{N}^{-1/2}(\alpha, \gamma, \lambda)$ . Figure 1 shows the three aforementioned cases, with parameters chosen such that three means are equal.

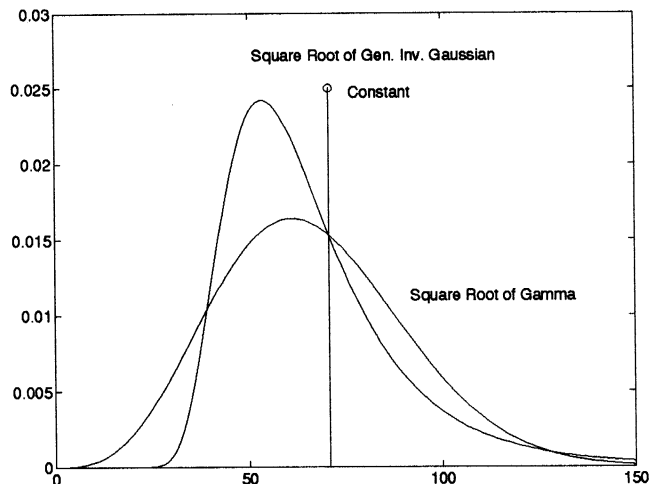


Figure 1: Amplitude backscatters: constant, square root of Gamma and square root of generalized inverse Gaussian densities.

#### 4 INTENSITY BACKSCATTER

The distributions associated to the intensity backscatter follow immediately applying the transformation  $X_I = X_A^2$  yielding, thus, to a constant  $X_I \sim \mathcal{C}(\beta)$  for the homogeneous case, and to the following densities corresponding to the heterogeneous and extremely heterogeneous cases (the densities are related by  $f_{X_A}(x) = 2x f_{X_I}(x^2)$ )

$$f_{X_I}(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x), \quad x, \alpha, \lambda > 0; \quad (3)$$

$$f_{X_I}(x) = \frac{(\lambda/\gamma)^{\alpha/2}}{2K_\alpha(2\sqrt{\lambda\gamma})} x^{\alpha-1} \exp\left(-\frac{\gamma}{x} - \lambda x\right), \quad (4)$$

with  $x > 0$  and the parameters space given in (1). The distributions characterized by densities (3) and (4) are called Gamma and generalized inverse Gaussian, resp., and will be here denoted as  $X_I \sim \Gamma(\alpha, \gamma)$  and  $X_I \sim \mathcal{N}^{-1}(\alpha, \gamma, \lambda)$ . Their  $r$ -th order moments are given, respectively, by

$$\begin{aligned} \mathbb{E}(X_I^r) &= \frac{\Gamma(\alpha+r)}{\lambda^r \Gamma(\alpha)} \\ \mathbb{E}(X_I^r) &= \left(\frac{\gamma}{\lambda}\right)^{r/2} \frac{K_{\alpha+r}(2\sqrt{\gamma\lambda})}{K_\alpha(2\sqrt{\gamma\lambda})}. \end{aligned}$$

Important limiting properties that relate the aforementioned distributions can be posed in the following manner

$$\mathcal{N}^{-1}(\alpha, \gamma, \lambda) \xrightarrow[\gamma \downarrow 0]{\mathcal{D}} \Gamma(\alpha, \lambda) \xrightarrow[\alpha/\lambda \rightarrow \beta]{\alpha, \lambda \rightarrow \infty} \mathcal{C}(\beta), \quad (5)$$

The properties stated in relations (5), in analogy with the amplitude case, assure the unicity of the model, since every type of intensity backscatter can be treated as the outcome of a special case of  $X_I \sim \mathcal{N}^{-1}(\alpha, \gamma, \lambda)$ . An extensive account of properties and applications of the generalized inverse Gaussian distribution can be seen in Jørgensen (1982). It should be noted that this distribution includes the distribution of Gamma ( $\gamma = 0$ ), reciprocal of Gamma ( $\lambda = 0$ ) and Hyperbola ( $\alpha = 1/2$ ) random variables as particular cases (Barndorff-Nielsen and Bæsild, 1981).

## 5 COMPLEX RETURN

Once defined the distributions for  $X_A$ , and given that the complex speckle is assumed to be  $\mathbf{Y}_C = (Y_{\Re}, Y_{\Im}) \sim \mathcal{N}(\mathbf{0}, \mathbf{1}/2)$ , it is immediate to derive the three associated marginal distributions for the complex return, which is given by  $\mathbf{Z}_C = X_A \cdot \mathbf{Y}_C = X_A \cdot (Y_{\Re}, Y_{\Im})$ ; they are  $\mathcal{N}(0, \beta/2)$  if the area is homogeneous and, for the heterogeneous and extremely heterogeneous cases respectively, the distributions characterized by the following densities

$$f_{Z_o}(x) = \frac{2}{\Gamma(\alpha)} \sqrt{\frac{\lambda^{\alpha+1/2}}{\pi}} |x|^{\alpha-1/2} K_{\alpha-1/2} \left( 2|x|\sqrt{\lambda} \right), \quad (6)$$

$$f_{Z_o}(x) = \frac{1}{K_{\alpha}(2\sqrt{\lambda\gamma})} \sqrt{\frac{(\lambda/\gamma)^{\alpha}}{\pi}} \left( \frac{\gamma + x^2}{\lambda} \right)^{(\alpha-1/2)/2} K_{\alpha-1/2} \left( 2\sqrt{\lambda(\gamma + x^2)} \right), \quad (7)$$

where  $Z_o$  denotes either the real or imaginary part of  $\mathbf{Z}_C$ . The distributions characterized by densities (6) and (7) are called, respectively, complex  $K$  and complex  $G$ , and are denoted here as  $\mathcal{K}_C(\alpha, \lambda)$  and  $\mathcal{G}_C(\alpha, \gamma, \lambda)$ , respectively. Their  $r$ -th order moments are given, respectively by

$$\begin{aligned} \mathbb{E}(Z_o^r) &= \begin{cases} 0 & \text{if } r \text{ is odd} \\ \frac{\Gamma(\alpha+r/2)}{(2\lambda)^{r/2}\Gamma(\alpha)} (1 \cdot 3 \cdot \dots \cdot (r-1)) & \text{else} \end{cases}, \\ \mathbb{E}(Z_o^r) &= \begin{cases} 0 & \text{if } r \text{ is odd} \\ \frac{K_{\alpha+r/2}(2\sqrt{\gamma\lambda})}{K_{\alpha}(2\sqrt{\gamma\lambda})} \left( \frac{\gamma}{4\lambda} \right)^{r/4} (1 \cdot 3 \cdot \dots \cdot (r-1)) & \text{else.} \end{cases} \end{aligned}$$

Important limiting properties that relate the aforementioned distributions can be posed in the following manner

$$\mathcal{G}_C(\alpha, \gamma, \lambda) \xrightarrow[\gamma \downarrow 0]{\mathcal{D}} \mathcal{K}_C(\alpha, \lambda) \xrightarrow[\alpha/\lambda \rightarrow \beta]{\mathcal{D}} \mathcal{N}(0, \beta/2), \quad (8)$$

The properties stated in relations (8), in analogy with the amplitude and intensity cases, assure the unicity of the model, since every type of complex return can be marginally treated as the outcome of a special case of  $Z_o \sim \mathcal{G}_C(\alpha, \gamma, \lambda)$ . Figure 2 shows the three aforementioned cases, with parameters chosen such that the variances of the three random variables are equal.

Using the parametrization presented here, it can be easily checked that the complex  $G$  distribution is a particular case of the generalized hyperbolic distribution. Many of its properties and applications can be seen in Barndorff-Nielsen and Bæsild (1981) and references therein.

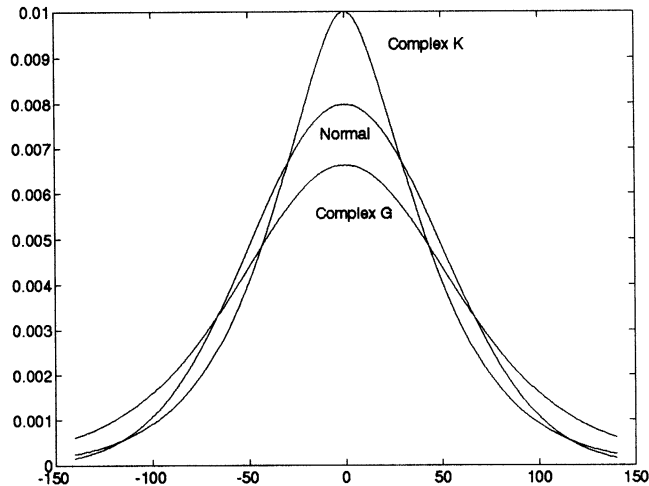


Figure 2: Complex returns: normal, complex  $K$  and complex  $G$  densities.

## 6 AMPLITUDE RETURN

In order to obtain the distribution of the amplitude return, the distribution of the product  $Z_A = X_A \cdot Y_A$  has to be computed for every case of the backscatter. It is possible to prove that from the constant backscatter follows a  $\sqrt{\Gamma}(n, n/\beta)$  distribution, with  $\beta = \alpha/\lambda$ , and that heterogeneous and extremely heterogeneous backscatters, as previously characterized, yield to distributions with the following densities

$$f_{Z_A}(x) = \frac{4\lambda n x}{\Gamma(\alpha)\Gamma(n)} (\lambda n x^2)^{(\alpha+n)/2-1} K_{\alpha-n} \left( 2x\sqrt{\lambda n} \right) \quad (9)$$

$$f_{Z_A}(x) = \frac{2n^n (\lambda/\gamma)^{\alpha/2}}{\Gamma(n)K_\alpha \left( 2\sqrt{\lambda/\gamma} \right)} x^{2n-1} \left( \frac{\gamma + n x^2}{\lambda} \right)^{(\alpha-n)/2} K_{\alpha-n} \left( 2\sqrt{\lambda(\gamma + n x^2)} \right) \quad (10)$$

The distributions characterized by the densities given in equations (9) and (10) are called amplitude multilook  $K$  and  $G$ , and here denoted by  $\mathcal{K}_A(\alpha, \lambda, n)$  and  $\mathcal{G}_A(\alpha, \gamma, \lambda, n)$ , respectively. Their  $r$ -th order moments are given, respectively, by

$$\begin{aligned} E(Z_A^r) &= (n\lambda)^{-r/2} \frac{\Gamma(n+r/2)\Gamma(\alpha+r/2)}{\Gamma(n)\Gamma(\alpha)}, \\ E(Z_A^r) &= \left( \frac{\gamma}{n^2\lambda} \right)^{r/4} \frac{K_{\alpha+r/2} \left( 2\sqrt{\gamma\lambda} \right) \Gamma(n+r/2)}{K_\alpha \left( 2\sqrt{\gamma\lambda} \right) \Gamma(n)}. \end{aligned}$$

These distributions are related by the following limiting properties

$$\mathcal{G}_A(\alpha, \gamma, \lambda, n) \xrightarrow[\gamma \downarrow 0]{\mathcal{D}} \mathcal{K}_A(\alpha, \lambda, n) \xrightarrow[\alpha/\lambda \rightarrow \beta]{\mathcal{D}} \sqrt{\Gamma}(n, n/\beta). \quad (11)$$

In this manner, every amplitude return can be seen as the outcome of a particular case of a  $\mathcal{G}_A(\alpha, \gamma, \lambda, n)$  distributed random variable. Figure 3 shows the three aforementioned cases, with the same parameters chosen for Figure 2 and  $n = 1$ .



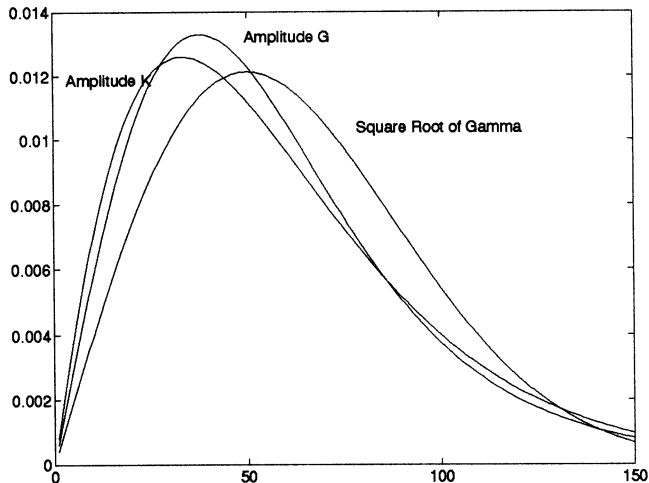


Figure 3: Amplitude returns: square root of Gamma, amplitude  $K$  and amplitude  $G$  densities.

## 7 INTENSITY RETURN

In order to obtain the distribution of the intensity return, the distribution of the product  $Z_I = X_I \cdot Y_I$  has to be computed for every case of backscatter, or simply to apply the aforementioned densities transformation  $f_{X_A}(x) = 2xf_{X_I}(x^2)$  to the amplitude densities. It is possible to prove that from the constant backscatter follows a  $\Gamma(\alpha, n/\beta)$  distribution, with  $\beta = \alpha/\lambda$ , and that heterogeneous and extremely heterogeneous backscatters, as previously characterized, yield to distributions characterized by the following densities

$$f_{Z_I}(x) = \frac{2\lambda n}{\Gamma(\alpha)\Gamma(n)} (\lambda nx)^{(\alpha+n)/2-1} K_{\alpha-n} \left( 2\sqrt{\lambda nx} \right) \quad (12)$$

$$f_{Z_I}(x) = \frac{n^n (\lambda/\gamma)^{\alpha/2}}{\Gamma(n)K_\alpha \left( 2\sqrt{\lambda/\gamma} \right)} x^{n-1} \left( \frac{\gamma + nx}{\lambda} \right)^{(\alpha-n)/2} K_{\alpha-n} \left( 2\sqrt{\lambda(\gamma + nx)} \right) \quad (13)$$

The distributions characterized by the densities given in equations (12) and (13) are called intensity multilook  $K$  and  $G$ , and here denoted by  $\mathcal{K}_I(\alpha, \lambda, n)$  and  $\mathcal{G}_I(\alpha, \gamma, \lambda, n)$ , respectively. These distributions are related by the following limiting properties

$$\mathcal{G}_I(\alpha, \gamma, \lambda, n) \xrightarrow[\gamma \downarrow 0]{\mathcal{D}} \mathcal{K}_I(\alpha, \lambda, n) \xrightarrow[\alpha/\lambda \rightarrow \beta]{\mathcal{D}_{\alpha, \lambda \rightarrow \infty}} \Gamma(n, n/\beta), \quad (14)$$

and their  $r$ -th order moments are given by  $E(Z_I^r) = E(Z_A^{2r})$ , where  $E(Z_A^r)$  is the  $r$ -th order moment of the corresponding amplitude random variable, which are given in Section 6.

In this manner, every intensity return can be seen as a particular case of the outcome of a  $\mathcal{G}_I(\alpha, \gamma, \lambda, n)$  distributed random variable. Figure 4 shows the three aforementioned cases, with the same parameters chosen for Figure 2 and  $n = 1$ .

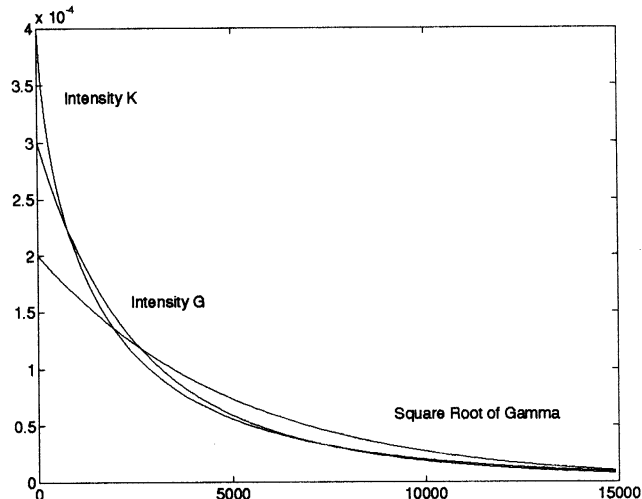


Figure 4: Intensity returns: Gamma, intensity  $K$  and intensity  $G$  densities.

In Yanasse et al. (1995) was presented the following recursive form for the cumulative distribution function for  $\mathcal{K}_I$  distributed random variables, where  $[x] = \max\{k \in \mathbb{N}: k \leq x + 1/2\}$

$$F_{Z_I}(x) = 1 + \frac{2^{2-\theta_1-\theta_2}}{\Gamma(\theta_1)\Gamma(\theta_2)}g(\nu, k, z),$$

where

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

and  $(\theta_1 = \alpha, \theta_2 = [n])$  or  $(\theta_2 = [\alpha], \theta_1 = n)$ ,  $\nu = \theta_1 - \theta_2$ ,  $k = 2\theta_2 - 1$  and  $z = \sqrt{\theta_1\theta_2\alpha x/\lambda}$ . The reader is referred to the previously mentioned report for details about the use of this expression. The cumulative distribution function of  $\mathcal{K}_A$  random variables is given by  $F_{Z_A}(x) = F_{Z_I}(x^2)$ . The cumulative distribution function of  $\Gamma(\alpha, \lambda)$  distributed random variables (and, therefore, of  $\sqrt{\Gamma}(\alpha, \lambda)$  ones) can be computed using standard numerical routines. The cumulative distribution function of  $\mathcal{G}$  distributed random variables, to the present date and to the knowledge of the authors, only can be computed by the numerical integration of the corresponding density.

## 8 NON-MULTIPLICATIVE DISTRIBUTIONS

In some special cases (Yanasse et al., 1995), even under the multiplicative model, the normal distribution could model the intensity and amplitude returns. Those cases arise in the following situations:

1. When the homogeneity is somewhere between the homogeneous and heterogeneous cases, and the number of looks is large enough ( $n \rightarrow \infty$ ) to assume that the speckle is a constant equal to 1. The intensity

backscatter, under these assumptions, has approximately a  $\mathcal{N}(\alpha/\lambda, \alpha/\lambda^2)$  distribution and, therefore, the return has the same distribution.

2. When there is complete homogeneity (and, therefore, the backscatter is a constant equal to  $\beta$ ) and the number of looks is large enough to assume that the speckle is  $\mathcal{N}(1, 1/n)$  distributed. Under these hypothesis the intensity backscatter has approximately a  $\mathcal{N}(\beta, \beta^2/n)$  distribution.

Distributions other than those arising from a formal application of the multiplicative model have also been successfully used in SAR image analysis. Some of them are commented here, and were used in the analysis reported in Yanasse et al. (1994). All these distributions are intended for the modelling of the intensity and amplitude returns (and, therefore, no subscript is used for the random variable  $Z$ ), without any explicit attempt to decompose it in backscatter and speckle.

The Log-Normal distribution is derived from  $V \sim \mathcal{N}(\mu, \sigma^2)$  and taking  $Z = \exp(V)$ . This case is denoted here as  $Z \sim \mathcal{LN}(\mu, \sigma^2)$ , and its density is given by

$$f_Z(z) = \frac{1}{z\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(z) - \mu}{\sigma}\right)^2\right), \quad \mu \in \mathbb{R}, \quad z, \sigma > 0.$$

The random variable  $Z$  is said to have Beta distribution with parameters  $\alpha, \beta > 0$ , denoted here as  $Z \sim \mathcal{B}(\alpha, \beta)$  if its density is given by

$$f_Z(z) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} z^{\alpha-1}(1-z)^{\beta-1}, \quad z \in (0, 1).$$

Obviously, when this distribution is intended to model image data, the density support has to be extended from the minimum to the maximum possible values.

The random variable  $Z$  is said to have Weibull distribution with parameters  $\alpha, \beta > 0$ , denoted here as  $Z \sim \mathcal{W}(\alpha, \beta)$  if its density is given by

$$f_Z(z) = \alpha\beta^\alpha z^{\alpha-1} \exp(-\beta z^\alpha), \quad z > 0.$$

Some works (Farina et al., 1987) report the use of a transformed Weibull random variable, namely if  $V \sim \mathcal{W}(\alpha, \beta)$ , then the data is modelled as  $Z = V^p$ , with  $p > 0$ .

## 9 APPLICATIONS

From the presented framework it can be deduced that, given a format and (if applicable) the equivalent number of looks, it is only necessary to estimate the parameters of the corresponding  $\mathcal{G}$  distribution since these values indicate the degree of homogeneity of the observed area, as seen in relations (8), (11) and (14). Distributions that do not fit into the multiplicative model as, for instance, those presented in Section 8 do not possess this nice discriminatory capability based on parameters.

As a first attempt to statistical discrimination of samples using the  $K$  distribution, in Yanasse et al. (1995) several samples from forest and non-forest were taken from a SAREX image over Tapajós, Brazil. This image is in amplitude format, and the equivalent number of looks was estimated as 5.6 (while the informed nominal

value is 7). Those samples were resampled in order to have approximately independent observations. Same samples were taken from the HH and VV polarizations.

The parameters of the square root of Gamma, amplitude  $K$  and Normal distributions were estimated for every sample, and the Kolmogorov-Smirnov test was applied to test the three hypothesis. The percentage of samples not rejected by this goodness-of-fit test at the 10% level is presented in Table 1. This table shows that the  $\mathcal{K}_A$  distribution also provides a good fit for both types of areas; this result was expected since this distribution is more general than the  $\sqrt{\Gamma}$  distribution, in the sense that the latter is a limiting case of the former. It is important to notice that the parameters of the  $\mathcal{K}_A$  distribution perform a discrimination between forest and non-forest samples. This discriminatory capability is shown in Fig. 5, where a clear distinction is drawn between forest and non-forest samples by  $\hat{\alpha}$ , the estimator of the parameter  $\alpha$  of the  $\mathcal{K}_A$  distribution. In this Figure it is noticeable that  $\hat{\alpha} \approx 20$  separates the two types of areas, suggesting that forest (non-forest) samples were generated by heterogeneous (homogeneous) backscatters since for these samples  $\hat{\alpha} < 20$  ( $\hat{\alpha} > 20$ , resp.)

Sample - Distribution	$\sqrt{\Gamma}$	$\mathcal{K}_A$	$\mathcal{N}$
Non-forest VV	87	87	100
Non-forest HH	100	97	100
Forest VV	88	95	39
Forest HH	92	88	38

Table 1: PERCENTAGE OF SAMPLES ACCEPTED AT THE 10% SIGNIFICANCE LEVEL.

It is also noticeable from Table 1 that the  $\mathcal{N}$  distribution fails to fit forest areas, as previewed by the theory. Some of the non-multiplicative distributions mentioned in Section 8 were also used to describe those samples and, though the results are interesting, they fail to explain the samples in terms of their homogeneity.

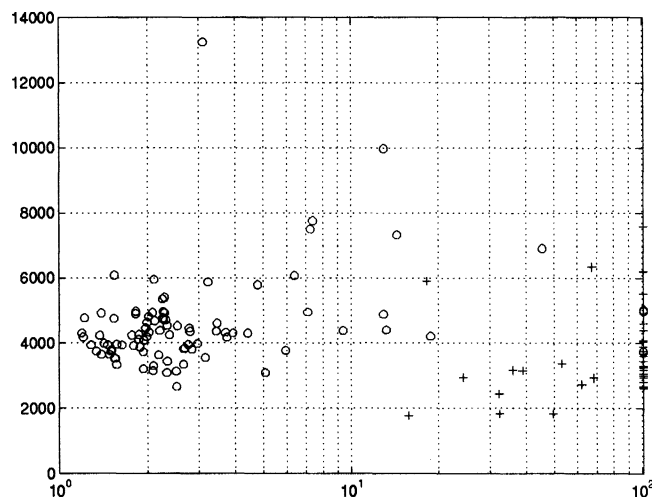


Figure 5: Estimated parameters  $\alpha$  and  $\beta$  of the  $\mathcal{K}_A$  distribution for forest (“o”) and non-forest (“+”) samples.

Two samples, one from non-forest (homogeneous area) and other from forest (heterogeneous area), from the VV polarization band were chosen to illustrate the adequacy of the presented technique. Figs. 6 and 7 show the histograms and fitted densities after parameters estimation of  $\mathcal{N}$ ,  $\mathcal{K}_A$  and  $\sqrt{\Gamma}$  distributions.

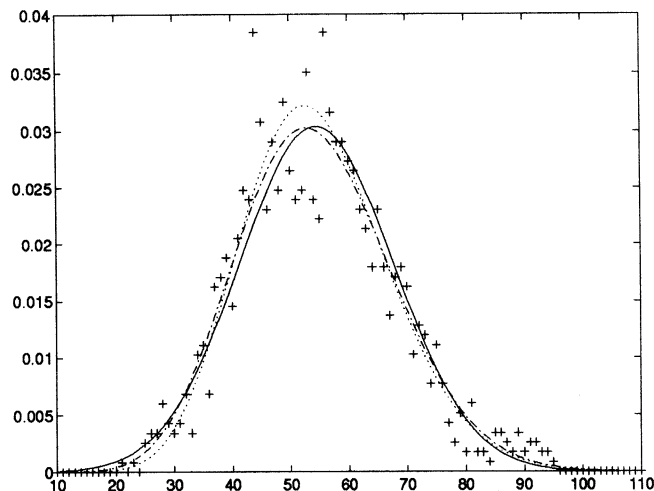


Figure 6: Histogram and fitted densities for a non-forest sample: normal (continuous), square root of Gamma (dot-dashed) and amplitude  $K$  (dotted lines, resp.)

The use of  $G$  distributions poses a delicate problem, namely parameter estimation, since the mathematical difficulties are remarkable. Some effort is being devoted to finding computationally suitable estimators for the parameters that characterize these distributions.

## 10 CONCLUSIONS AND EXTENSIONS

A unified use of the multiplicative model was presented to derive some useful distributions for SAR image modelling and analysis, for several degrees of heterogeneity and three usual image formats. Many relationships relating these distributions are presented. The cumulative distribution functions of multilook  $K$  distributed random variables was provided in a recursive form, that requires the discretization of one of its three parameters. An extension of the “classic” distributions that arise within the multiplicative model is achieved through the use of the square root of generalized inverse Gaussian distribution for the amplitude backscatter, giving rise to the class of  $\mathcal{G}$  distributions for the return. The estimation of the parameter  $\alpha$  of  $K$  distributions was seen as a good discriminatory technique.

Multivariate distributions could also be considered to exploit the maximum of information from polarimetric data. It is interesting to note that the generalized hyperbolic distribution (Bardoff-Nielsen and Blæsild, 1981) is defined for an arbitrary number of dimensions, and so is the  $\mathcal{G}_C$  distribution. Efficient estimation techniques for  $G$  distributions is an open problem.

The knowledge of the statistical properties of SAR data allows the proposal of specialized processing/segmentation/classification techniques, better suited for dealing with the speckle noise than those used for optical data. For instance, in Frery and Sant’Anna (1993) some specialized filters are proposed for the Rayleigh distribution, which is the adequate model for homogeneous areas, amplitude detection and one look. Those *robust* filters

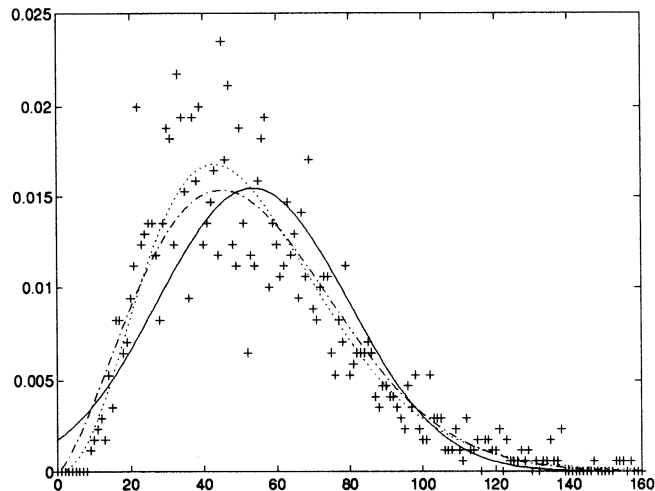


Figure 7: Histogram and fitted densities for a forest sample: normal (continuous), square root of Gamma (dot-dashed) and amplitude  $K$  (dotted lines, resp.)

produce a remarkable speckle noise reduction, while preserving edges and sharp features, and could be extended to other formats and heterogeneity cases.

#### ACKNOWLEDGEMENTS

The authors are grateful to FAPESP (Project 91/3532-2), and to CNPq (Project GEOTEC 680.061-94-0).

#### REFERENCES

- Barndorff-Nielsen, O.E.; Blæsild, P. Hyperbolic distributions and ramifications: contributions to theory and applications. In: Taillie, C. and Baldessari, B.A. (eds.) *Statistical distributions in scientific work*. Reidel, Dordrecht, 1981, p. 19–44.
- Caves, R.G. *Automatic matching of features in synthetic aperture radar to digital map data*. (PhD Thesis) — University of Sheffield, Sheffield, UK, 1993.
- Farina, A.; Russo, A.; Scannapieco, F.; Barbarossa, S. Theory of radar detection in coherent Weibull clutter. *Proc. IEE*, **134**(F):174–190, 1987.
- Frery, A.C.; Sant’Anna, S.J.S. Non-adaptive robust filters for speckle noise reduction. In: Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens, 6., Recife, PE, 19–22 out. 1993. *Anais*. Recife, SBC/UFPe, out. 1993, p. 165–174.

- Jakeman, E. On the statistics of K-distributed noise. *Journal of Physics A: Mathematical and General*, **6L**:89–92, 1973.
- Jakeman, E.; Pusey, P.N. A model for non-Rayleigh sea echo. *IEEE Trans. on Antennas and Propagation*, **AP-24**(6):806–814, 1976.
- Jao, J.K. Amplitude distribution of composite terrain radar clutter and the K distribution. *IEEE Trans. on Antennas and Propagation*, **AP-32**:1049–1061, 1984.
- Jørgensen, B. *Statistical properties of the generalized inverse Gaussian distribution*. New York, Springer-Verlag, 1982. (Lecture Notes in Statistics, 9)
- Yanasse, C.C.F.; Frery, A.C.; Sant'Anna, S.J.S.; Hernandez Filho, P.; Dutra, L.V. Statistical analysis of SAREX data over Tapajós – Brazil. In: SAREX-92 South American Radar Experiment, Paris, 6–8 Dec. 1992. *Workshop Proceedings*. Paris, ESA, Mar. 1994, p. 25–40.
- Yanasse, C.C.F.; Frery, A.C.; Sant'Anna, S.J.S. *Stochastic distributions and the multiplicative model: relations, properties, estimators and applications to SAR image analysis*. INPE, São José dos Campos, 1995. (INPE-5630-NTC/318).