MIMICK CAPACITY OF GENERALIZED GAMMA DISTRIBUTION FOR HIGH RESOLUTION SAR IMAGE STATISTICAL MODELING

Helene Sportouche, Jean-Marie Nicolas, Florence Tupin

LTCI, CNRS, Telecom ParisTech, Universite Paris-Saclay, 75013, Paris, France

ABSTRACT

In this paper we investigate the capacity of the Generalized Gamma distribution to mimick (or imitate) thanks to its three parameters other useful SAR distributions. We first compare it with the Fisher distribution when mimicking a $\mathcal{K}$ distribution of reference, thanks to the log-cumulant approach and through a Kullback-Leibler divergence. We then study how the Generalized Gamma distribution can imitate a Log-Normal distribution as asymtotic limit.

***Index Terms*—** Statistical modeling of SAR data, Generalized Gamma distribution, log-cumulant parameter estimation

1. INTRODUCTION

The choice of an appropriate distribution for the statistical modeling of SAR data is essential for many applications like despeckling, segmentation or classification [1] [2]. Beyond physical modeling of electro-magnetic wave interactions [3][4], many complex distributions have already been proposed in the literature to model high resolution SAR images. Specially, the Generalized Gamma (GG) distribution [5] (defined by three parameters) has been the subject of many studies and its flexibility is demonstrated both for physically homogeneous or heterogeous areas [6]. Nevertheless is it general enough to replace a distribution dictionary ?

In this paper we investigate the capacity of the Generalized Gamma distribution to imitate a reference distribution. We first choose the $\mathcal{K}$ [4] distribution also defined by 3 parameters and we compare the GG mimicking capacity with the mimicking capacity of the Fisher distribution [7]. In this study we use a log-cumulant approach to estimate the distribution parameters and use evalutation criteria based on Kullback-Leibler divergence.

We then discuss some limits of the Generalized Gamma distribution and study its relationship with the Log-Normal distribution [8] (defined by two parameters).

The paper is organized as follows. In section 2, analytic expressions of the distributions are recalled while section 3 presents the comparative study of $\mathcal{K}$ distribution mimicking. Eventually section 4 presents how Generalized Gamma can be used to mimick Log-Normal distributions.

Thanks to ANR ALYS funding.

2. PRELIMINARIES

2.1. Distribution presentation

We remind here the probability density function (pdf) of the Generalized Gamma distribution and its associated log-cumulants. For Fisher, $\mathcal{K}$ and Log-Normal [9] only the log-cumulants are presented.

2.1.1. Generalized Gamma distribution

The pdf $\mathcal{GG}_{pdf}(\mu_{\mathcal{GG}}, L_{\mathcal{GG}}, \eta_{\mathcal{GG}})$ is defined by:

$$\begin{aligned}\mathcal{GG}_{pdf}(x) \quad = \quad & \frac{|\eta_{\mathcal{GG}}|}{\mu_{\mathcal{GG}}} \frac{L_{\mathcal{GG}}^{(\frac{1}{\eta_{\mathcal{GG}}})}}{\Gamma(L_{\mathcal{GG}})} \left(\frac{L_{\mathcal{GG}}^{(\frac{1}{\eta_{\mathcal{GG}}})} x}{\mu_{\mathcal{GG}}} \right)^{\eta_{\mathcal{GG}} L_{\mathcal{GG}} - 1} \\ & e^{-\left(\frac{L_{\mathcal{GG}}^{(\frac{1}{\eta_{\mathcal{GG}}})} x}{\mu_{\mathcal{GG}}} \right)^{\eta_{\mathcal{GG}}}}\end{aligned} \tag{1}$$

where $\Gamma(x)$is the Gamma function, $\mu_{\mathcal{GG}}$ a mean parameter, $L_{\mathcal{GG}}$ a shape parameter and $\eta_{\mathcal{GG}}$ a power parameter.

The 3 first log-cumulants are given by:

$$\begin{aligned}\widetilde{\kappa_1}_{\mathcal{GG}_{pdf}} &= \ln(\mu_{\mathcal{GG}}) + \frac{\Psi(L_{\mathcal{GG}}) - \ln(L_{\mathcal{GG}})}{\eta_{\mathcal{GG}}} \\ \widetilde{\kappa_2}_{\mathcal{GG}_{pdf}} &= \frac{\Psi(1, L_{\mathcal{GG}})}{\eta_{\mathcal{GG}}^2} \\ \widetilde{\kappa_3}_{\mathcal{GG}_{pdf}} &= \frac{\Psi(2, L_{\mathcal{GG}})}{\eta_{\mathcal{GG}}^3}\end{aligned} \tag{2}$$

where $\Psi(x)$ and $\Psi(n,x)$ are respectively the Digamma and the Polygamma function of order n.

2.1.2. Fisher distribution

The 3 first log-cumulants are given by:

$$\begin{aligned}\widetilde{\kappa_1}_{\mathcal{F}_{pdf}} &= \ln(\mu_{\mathcal{F}}) + \Psi(L_{\mathcal{F}}) - \ln(L_{\mathcal{F}}) - \Psi(M_{\mathcal{F}}) + \ln(M_{\mathcal{F}}) \\ \widetilde{\kappa_2}_{\mathcal{F}_{pdf}} &= \Psi(1, L_{\mathcal{F}}) + \Psi(1, M_{\mathcal{F}}) \\ \widetilde{\kappa_3}_{\mathcal{F}_{pdf}} &= \Psi(2, L_{\mathcal{F}}) - \Psi(2, M_{\mathcal{F}})\end{aligned} \tag{3}$$

where $\mu_{\mathcal{F}}$ is a mean parameter, $L_{\mathcal{F}}$ and $M_{\mathcal{F}}$ are both shape parameters.

2.1.3. K distribution

The 3 first log-cumulants are given by:

$$\begin{array}{rcl}\widetilde{\kappa_1}_{\mathcal{K}_{pdf}} & = & \ln(\mu_{\mathcal{K}}) + \Psi(L_{\mathcal{K}}) - \ln(L_{\mathcal{K}}) + \Psi(M_{\mathcal{K}}) - \ln(M_{\mathcal{K}}) \\ \widetilde{\kappa_2}_{\mathcal{K}_{pdf}} & = & \Psi(1, L_{\mathcal{K}}) + \Psi(1, M_{\mathcal{K}}) \\ \widetilde{\kappa_3}_{\mathcal{K}_{pdf}} & = & \Psi(2, L_{\mathcal{K}}) + \Psi(2, M_{\mathcal{K}}) \end{array} \tag{4}$$

where $\mu_{\mathcal{K}}$ is a mean parameter, $L_{\mathcal{K}}$ and $M_{\mathcal{K}}$ are two shape parameters. The special case of $\mathcal{K}$ distribution where $L_{\mathcal{K}} = M_{\mathcal{K}}$ corresponds to *caustic* $\mathcal{K}_c$ distributions.

2.1.4. Log-Normal distribution

The 3 first log-cumulants are given by:

$$\begin{array}{rcl}\widetilde{\kappa_1}_{\mathcal{LN}_{pdf}} & = & \mu_{\mathcal{LN}} \\ \widetilde{\kappa_2}_{\mathcal{LN}_{pdf}} & = & {\sigma_{\mathcal{LN}}}^2 \\ \widetilde{\kappa_3}_{\mathcal{LN}_{pdf}} & = & 0 \end{array} \tag{5}$$

where $\mu_{\mathcal{LN}}$ is a mean parameter and $\sigma_{\mathcal{LN}}$ a standard deviation parameter. The log-cumulants of the Log-Normal are null for all order above 2.

2.2. Representation in the log-cumulant diagram of order 2 and 3

We present in figure 1 the positionning of the distributions (regions for 3-parameters pdf and lines for 2-parameters pdf). The large areas covered by Generalized Gamma distributions and Fisher distributions explain their generic shape (please, note that the ordinate axis is excluded from the area covered by GG pdf). Both GG pdf and Fisher pdf areas contain the $\mathcal{K}$ distributions.

To apply an estimation method of the parameters based on log-cumulants (denoted by *LogCum*) it is necessary that the following conditions are fulfilled [10]:

$$\frac{\widehat{\kappa_2}^3_{\mathcal{GG}_{pdf}}}{\widehat{\kappa_3}^2_{\mathcal{GG}_{pdf}}} > \frac{1}{4} \tag{6}$$

$$\widehat{\kappa_2}_{\mathcal{F}_{pdf}} \geq \Psi\left(1, \Phi(2, -\left|\widehat{\kappa_3}_{\mathcal{F}_{pdf}}\right|)\right) \tag{7}$$

where $\Phi(n, x)$ represents the inverse of the Polygamma of order n.

3. COMPARISON OF FISHER AND GG PDFS MIMICKING $\mathcal{K}$ DISTRIBUTIONS

3.1. Methodology

3.1.1. LogCum parameter estimation for mimicking pdf

Let us suppose that the $\mathcal{K}$ pdf parameters to be mimicked are unknown (as it is the case in practice). Only samples following the $\mathcal{K}$ pdf are available (N_s intensity values).

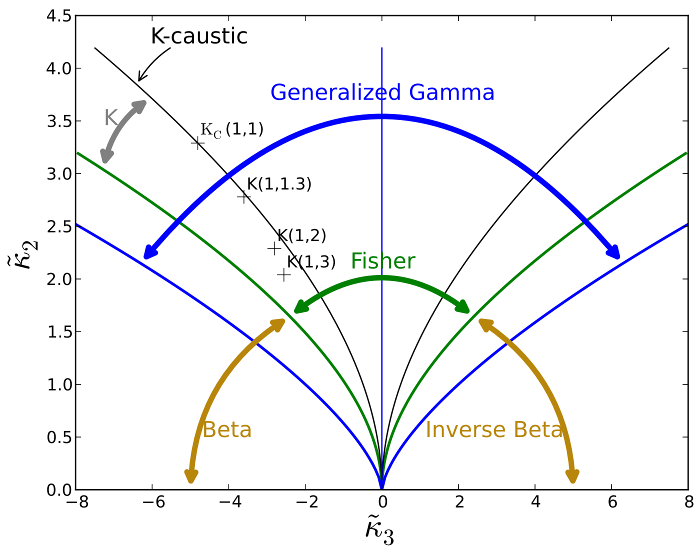

Fig. 1. Pdf positioning in the $\widetilde{\kappa_2} - \widetilde{\kappa_3}$ diagram: Generalized Gamma area (between blue curves excluding ordinate axis), Fisher area (between gree curves), $\mathcal{K}$ area (in grey) and Log-Normal distributions along the ordinate axis.

Parameter estimation of the mimicking distribution (GG or Fisher) is realized through the 3 first log-cumulants of the $\mathcal{K}$ pdf:

1. Computation of the empirical log-cumulants $\widehat{\kappa_2}$, $\widehat{\kappa_3}$ then $\widehat{\kappa_1}$ from the samples (see [9]).

2. Identification of these values using analytical expressions of $\widetilde{\kappa_2}_{\mathcal{GG}_{pdf}}$, $\widetilde{\kappa_3}_{\mathcal{GG}_{pdf}}$ then $\widetilde{\kappa_1}_{\mathcal{GG}_{pdf}}$ (or $\widetilde{\kappa_2}_{\mathcal{F}_{pdf}}$, $\widetilde{\kappa_3}_{\mathcal{F}_{pdf}}$ then $\widetilde{\kappa_1}_{\mathcal{F}_{pdf}}$), to estimate $L_{\mathcal{GG}}$, $\eta_{\mathcal{GG}}$ then $\mu_{\mathcal{GG}}$ or $L_{\mathcal{F}}$, $M_{\mathcal{F}}$ then $\mu_{\mathcal{F}}$ of the mimicking pdf.

This procedure will be applied in the experimental part to compare the mimicking capacity of GG and Fisher distributions.

3.1.2. Kullback-Leibler divergence (KLD) for pdf comparison

KLD measures the approximation of $\mathcal{P}$ (observation) by $\mathcal{Q}$ (model) :

$$KLD(\mathcal{P}_{pdf}(x), \mathcal{Q}_{pdf}(x)) = \int_0^\infty \mathcal{P}_{pdf}(x) \ln\left(\frac{\mathcal{P}_{pdf}(x)}{\mathcal{Q}_{pdf}(x)}\right) \mathrm{d}x \tag{8}$$

We propose to use the mean and standard-deviation of the KLD values computed between the $\mathcal{K}$ pdf and GG pdf (or Fisher pdf) with the following operating procedure:

1. Definition (i) of an initial parameter set for the $\mathcal{K}$ pdf and (ii) a sample size N_s
2. Generation of many realizations drawn from $\mathcal{K}$
3. For each test sample, (i) estimation of the mimicking pdf parameters (GG or Fisher) following 3.1.1 and (ii) computation of the KLD between $\mathcal{K}$ and the estimated distribution
4. Computation of KLD mean (for the *accuracy* study) and the KLD variance (for the *robustness* study).

The smaller the KLD mean the more accurate is the mimick, the smaller the KLD variance the more robust it is.

3.2. Results

Different situations have been studied to compare the mimicking capacities of GG and Fisher distributions. They are presented in figure 1, where crosses indicate the 4 processed test packs leading to similar conclusions. KLD means and sdandard deviations have been computed using 50 realizations and the test sample size N_s was chosen between 250 et 12 500.

We observe that GG pdf is slightly better than Fisher for accuracy (KLD means) whatever the sample size, and slightly better for robustness when the sample size N_s is bigger than a given sample size (500 pixels for the example presented in figure 2).

Supposing known the $\mathcal{K}$ pdf parameters to be mimicked, it is possible to compute the KLD values with the mimicking pdf (using the "true" derived parameters instead of the estimated ones). These are the convergence values of the KLD means when N_s tends to infinity (indicated in full line on the figure 2).

A visual analysis of the pdf shapes by superimposition of the curves on a same figure confirms that both distributions are very accurate to mimick another one. Again, a slight advantage for the Generalized Gamma distribution compared to Fisher for $\mathcal{K}$ mimicking is observed.

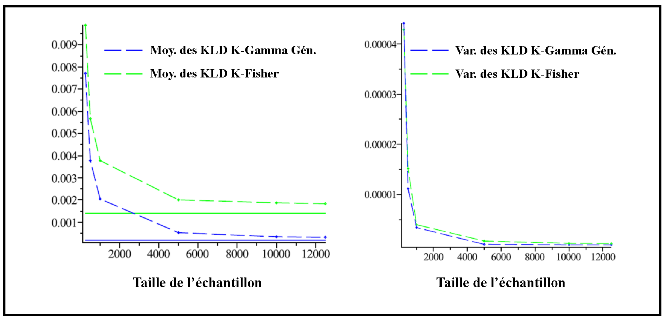

Fig. 2. Evolution of KLD means and variances related to the N_s sample size for $\mathcal{K}$ mimicking by GG and Fisher pdfs for the test pack $\mathcal{K}(1, 1, 3)$.

4. LIMITS OF THE GG PDF AND RELATIONSHIP WITH LOG-NORMAL PDF

4.1. Practical and theoretical comments

When using a GG pdf for SAR data modeling, the following considerations should be taken into account:

(1) The use of LogCum method should be verified (see equation 6) for GG parameter estimation (see [6] and [10] for practical cases).

(2) GG is theoretically not defined for ordinate axis so it should be checked that among the couples $(\widehat{\kappa_2}_{\mathcal{GG}_{pdf}}, \widehat{\kappa_3}_{\mathcal{GG}_{pdf}})$ in the log-cum diagram, none of them is situated along the Log-Normal axis.

(3) Numerical instabilities during Polygamma inversion should be kept in mind in some critical areas: along the ordinate axis $\eta_{\mathcal{GG}}$ parameter tends to zero.

In the following section we propose an alternative solution to some of these limits.

4.2. Log-Normal pdf as asymtotic limit of GG

In this section it is shown that GG can approximate a Log-Normal pdf when using an original parameter estimation method for the mimicking distribution.

By expressing $L_{\mathcal{GG}}$ as $L_{\mathcal{GG}} = \frac{S_{\mathcal{GG}}}{\eta^2_{\mathcal{GG}}}$, it can be demonstrated using $\Psi(r, x) \sim \frac{\Gamma(r)}{x^r}$, for x tending to $+\infty$, that:

$$\lim_{\eta_{\mathcal{GG}} \to 0} \widetilde{\kappa_2}_{\mathcal{GG}_{pdf}} = \lim_{\eta_{\mathcal{GG}} \to 0} \frac{\Psi(1, \frac{S_{\mathcal{GG}}}{\eta^2_{\mathcal{GG}}})}{\eta^2_{\mathcal{GG}}} = \frac{1}{S_{\mathcal{GG}}} \qquad (9)$$

and all the log-cumulants $\widetilde{\kappa_r}_{\mathcal{GG}_{pdf}}$ of order r strictly superior than 2 are null, when $\eta_{\mathcal{GG}}$ tends towards zero.

By identifying the order 1 and 2 log-cumulants of the GG pdf with those of the Log-Normal, we obtain that $\frac{1}{S_{\mathcal{GG}}} = {\sigma_{\mathcal{LN}}}^2$ and $\ln(\mu_{\mathcal{GG}}) = \mu_{\mathcal{LN}}$ when $\eta_{\mathcal{GG}}$ tends to 0, thanks to an approximation of the Digamma function $\Psi(x) \sim \log(x)$, for x tending to $+\infty$.

It is thus shown that the Log-Normal pdf can be defined as the limit of a GG pdf. The GG pdf is converted in a 2 parameters pdf by imposing a specified expression for $L_{\mathcal{GG}}$. Complementary mathematical justifications of the limit when $\eta_{\mathcal{GG}}$ tends to 0 will be given in the final paper.

Examples of GG pdf mimicking Log-Normal pdf are presented in figure 3 in the $\widetilde{\kappa_2} - \widetilde{\kappa_3}$ diagram.

The flatness of the curves around the ordinate axis is related to the numerical stability of the approximation.

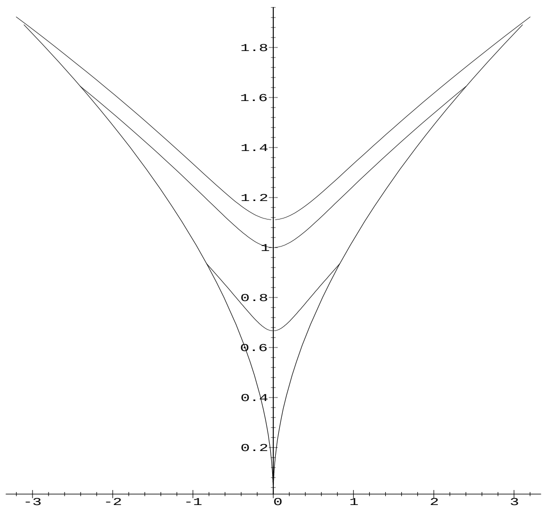

Fig. 3. Mimicks of the Log-Normal pdf by GG pdf for $\eta_{\mathcal{GG}}$ in $[-1; 1]$ and for $S_{\mathcal{GG}} \in \{0.9; 1; 1.5\}$ in the log-cumulant diagram. The curves representing Gamma and Inverse Gamma distributions are drawn on both sides of the mimicks for values $L_{\mathcal{G}} \in [0.9, \infty[$ and $L_{\mathcal{GI}} \in [0.9, \infty[$.

5. CONCLUSIONS AND FURTHER WORK

In a first part, we have observed that Generalized Gamma pdf is slightly more powerful for mimicking capacity than Fisher through the study of a mimicked reference $\mathcal{K}$ pdf. In a second part, it has been demonstrated that it is possible to approximate Log-Normal pdf through GG pdf by choosing a specific parametrization.

Further work includes a more general study of the mimicking capacity and an experimental validation using real samples of high resolution SAR images.

6. REFERENCES

[1] Alejandro César Frery, Hans-Jurgen Muller, Corina da Costa Freitas Yanasse, and Sidnei João Siqueira Sant'Anna, "A model for extremely heterogeneous clutter," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 3, pp. 648–659, 1997.

[2] Frédéric Galland, Jean-Marie Nicolas, Hélène Sportouche, Muriel Roche, Florence Tupin, and Philippe Réfrégier, "Unsupervised synthetic aperture radar image segmentation using Fisher distributions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 8, pp. 2966–2972, 2009.

[3] Joseph W Goodman, "Some fundamental properties of speckle," *Journal of the Optical Society of America*, vol. 66, no. 11, pp. 1145–1150, 1976.

[4] E. Jakeman and P. N. Pusey, "A model for non-Rayleigh sea echo," *IEEE TAP*, vol. 24, no. 6, pp. 806–814, 1976.

[5] E W Stacy, "A Generalization of the Gamma distribution," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1187–1192, 1962.

[6] Heng-Chao Li, Wen Hong, Yi-Rong Wu, and Ping-Zhi Fan, "On the empirical-statistical modeling of SAR images with generalized Gamma distribution," *IEEE J-STSP*, vol. 5, no. 3, pp. 386–397, 2011.

[7] Céline Tison, Jean-Marie Nicolas, Florence Tupin, and Henri Maître, "A new statistical model for markovian classification of urban areas in high-resolution SAR images," *IEEE TGRS*, vol. 42, no. 10, pp. 2046–2057, 2004.

[8] G. V. Trunk and S. F. George, "Detection of targets in non-Gaussian sea clutter," *IEEE TAES*, , no. 5, pp. 620–628, 1970.

[9] Jean-Marie Nicolas, "Application de la transformée de Mellin : étude des lois statistiques de l'imagerie cohérente," Tech. Rep. 2006D010, Telecom ParisTech, Paris, France, 2006.

[10] V. A. Krylov, G. Moser, S. B. Serpico, and J. Zerubia, "On the method of logarithmic cumulants for parametric probability density function estimation," *IEEE TIP*, vol. 22, no. 10, pp. 3791–3806, 2013.