Comparing samples from the $\mathcal{G}^0$ distribution using a geodesic distance

Alejandro C. Frery[1] · **Juliana Gambini**[2,3]

Abstract

The $\mathcal{G}^0$ distribution is widely used for monopolarized SAR image modeling because it can characterize regions with different degrees of texture accurately. It is indexed by three parameters: the number of looks (which can be estimated for the whole image), a scale parameter and a texture parameter. This paper presents a new proposal for comparing samples from the $\mathcal{G}^0$ distribution using a geodesic distance (GD) as a measure of dissimilarity between models. The objective is quantifying the difference between pairs of samples from SAR data using both local parameters (scale and texture) of the $\mathcal{G}^0$ distribution. We propose three tests based on the GD which combine the tests presented in Naranjo-Torres et al. (IEEE J Sel Top Appl Earth Obs Remote Sens 10(3):987–997, 2017), and we estimate their probability distributions using permutation methods.

Keywords Geodesic distance · Dissimilarity measure · $\mathcal{G}^0$ distribution

Mathematics Subject Classification 62F10 Point estimation · 62F03 Hypothesis testing · 62F40 Bootstrap, jackknife and other resampling methods · 62G32 Statistics of extreme values; tail inference

✉ Alejandro C. Frery
acfrery@gmail.com; acfrery@laccan.ufal.br

Juliana Gambini
juliana.gambini@gmail.com; mgambini@itba.edu.ar

1 Laboratório de Computação Científica e Análise Numérica, Universidade Federal de Alagoas, Av. Lourival Melo Mota, s/n, Maceió, AL 57072-900, Brazil

2 Depto. de Ingeniería Informática, Instituto Tecnológico de Buenos Aires, Av. Madero 399, C1106ACD Buenos Aires, Argentina

3 Depto. de Ingeniería en Computación, Universidad Nacional de Tres de Febrero, Sáenz Peña, Pcia. de Buenos Aires, Argentina

1 Introduction

Automatic detection of differences between samples from *Synthetic Aperture Radar* (SAR) images is both challenging and necessary. It has important applications in, among others, urban planning (Storie et al. 2012), disaster management (Dell'Acqua and Gamba 2012), emergency response (Sun et al. 2016), environmental monitoring and ecology (Hill et al. 2005). The main idea is developing methods for automatic discrimination of regions with different levels of texture and/or roughness. As in Gambini et al. (2006, 2008), we adopt the $\mathcal{G}^0$ distribution as model for the data.

The $\mathcal{G}^0$ distribution is widely used for monopolarized SAR image modeling because it can characterize different regions accurately. It is indexed by three parameters: the number of looks L (which can be estimated for the whole image), a scale parameter γ and a texture parameter α. The last two are local parameters and relate directly to the target.

Nascimento et al. (2010) obtained test statistics based on information theory to assess the null hypothesis that two samples were produced by the same $\mathcal{G}^0$ law, provided the same number of looks is known. The approach consisted of first computing $h - \phi$ divergences between the models, indexing their symmetrized versions with maximum likelihood estimates and scaling appropriately to obtain test statistics. These tests, under mild regularity conditions, follow asymptotically χ^2 laws. These divergences and associated test statistics were successfully applied to region discrimination (Silva et al. 2013), segmentation (Marques et al. 2012) and parameter estimation (Gambini et al. 2015).

Two issues make their use somewhat difficult, though, namely (1) they require the numerical integration of expressions that, more often than not, involve special functions, and (2) the choice of the particular test statistic might be considered arbitrary (different choices of the functions h and ϕ lead, among infinitely many others, to the Kullback–Leibler, Hellinger, Bhattacharya, Triangular, Harmonic, Jensen–Shannon, and Rényi of order β divergences). The geodesic distance solves the second difficulty, as it is unique, and gives a partial solution to the first one.

The geodesic distance can be used to measure the difference between two parametric distributions. It was presented by Rao (1945, 1992), and since then, it has been studied by several authors Verdoolaege and Scheunders (2012), Menendez et al. (1995) and Atkinson and Mitchell (1981). In Verdoolaege and Scheunders (2011) and Ilea et al. (2015), it is used as measure of contrast between samples by means of statistical tests presented in Salicrú et al. (1994), Menendez et al. (1995) and Nascimento et al. (2010), where the authors demonstrated that its distribution is χ^2_1.

To the best of the authors' knowledge, there is no closed expression for the geodesic distance between two $\mathcal{G}^0$ models with both α and γ unknown, given L. In this work, we analyze several statistical hypothesis tests depending on both parameters to discriminate two samples from $\mathcal{G}^0$ models with both parameters unknown. We use permutation methods to estimate the distribution of such tests statistics since no explicit results are available.

The paper unfolds as follows: Sect. 2 recalls properties of the $\mathcal{G}^0$ model, including parameter estimation by maximum likelihood. Section 3 presents the expressions for the GD with one parameter known. Section 4 analyzes the behavior of the test

statistics based on a known parameter. In Sect. 5, we study the more realistic situation of estimating both scale and texture, while assuming known the number of looks. Finally, in Sect. 6 we present conclusions and outline future work.

2 SAR imagery and the $\mathcal{G}^0$ model

Under the multiplicative model, the return in monopolarized SAR images can be modeled as the product of two independent random variables, one corresponding to the backscatter X and other to the speckle noise Y. In this manner, $Z = XY$ models the return Z in each pixel. For monopolarized data, speckle Y is modeled as a Γ distributed random variable with unitary mean and shape parameter L, the number of looks. A good choice for the backscatter distribution X is the reciprocal of Gamma $\Gamma^{-1}(\alpha, \gamma)$ law that gives rise to the $\mathcal{G}^0$ distribution for the return Z (Frery et al. 1997). The mathematical tractability and descriptive power of the $\mathcal{G}^0$ distribution make it an attractive choice for SAR data modeling (Quartulli and Datcu 2004). The probability density function for intensity data under the $\mathcal{G}^0(\alpha, \gamma, L)$ distribution is:

$$f_{\mathcal{G}^0}(z) = \frac{L^L \Gamma(L-\alpha)}{\gamma^\alpha \Gamma(-\alpha)\Gamma(L)} \frac{z^{L-1}}{(\gamma + zL)^{L-\alpha}}, \tag{1}$$

where $-\alpha, \gamma, z > 0$ and $L \geq 1$. If $\alpha \to -\infty$, the $\mathcal{G}^0$ distribution becomes an exponential law. The r-order moments are given by

$$E(Z^r) = \left(\frac{\gamma}{L}\right)^r \frac{\Gamma(-\alpha - r)}{\Gamma(-\alpha)} \frac{\Gamma(L + r)}{\Gamma(L)}. \tag{2}$$

To simplify calculation and with the intention of obtaining comparable results, in most experiments, we deal with a restricted case which assumes $E(Z) = 1$.

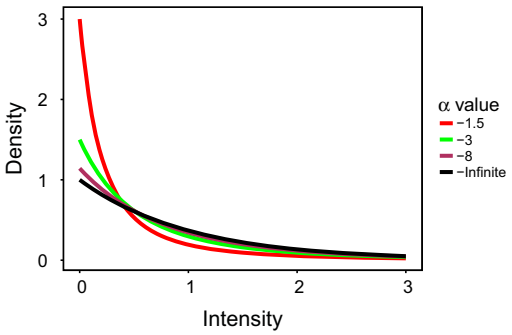
Using that $\Gamma(L + 1) = L\Gamma(L)$ and that $\Gamma(-\alpha) = (-\alpha - 1)\Gamma(-\alpha - 1)$ in (2), assuming $L = 1$ and imposing $E(Z) = 1$, we find the following relation between α and γ:

$$\gamma^* = -\alpha - 1.$$

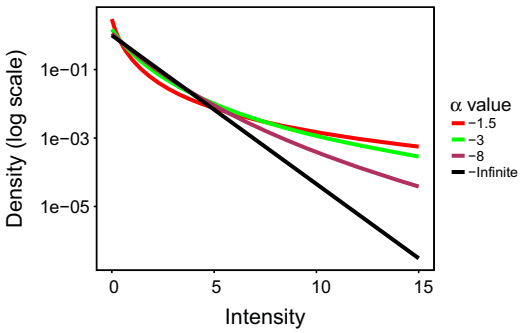
Then, the random variable Z with $\mathcal{G}_I^0(\alpha, \gamma^*, 1)$ distribution has unitary mean. This allows us to simplify the calculations and to obtain results which do not depend on image brightness.

One of the essential features of the $\mathcal{G}^0$ distribution is the ability to interpret its parameters. The α parameter is a texture parameter, which is related to the roughness or number of elementary backscatterers of the target. Values close to zero (typically above -3) suggest extremely textured targets, as urban zones. As the value decreases, it indicates regions with moderate texture (usually $\alpha \in [-6, -3]$), as forest zones. Textureless targets, e.g., pasture, usually produce $\alpha \in (-\infty, -6)$.

The γ parameter of the $\mathcal{G}^0$ distribution is a scale parameter, that is, if $W \sim \mathcal{G}^0(\alpha, \gamma, L)$, then $\gamma^{-1} W \sim \mathcal{G}^0(\alpha, 1, L)$.



(a) Densities in linear scale



(b) Densities in semilogarithmic scale

Fig. 1 Densities of $\mathcal{G}^0(\alpha, \gamma^*, 1)$ distributions for $\alpha \in \{-\infty, -8, -3, -1.5\}$ (black, maroon, green, red, respectively) (color figure online)

Figure 1 shows the densities of $\mathcal{G}^0(\alpha, \gamma^*, 1)$ distributions for $\alpha \in \{-\infty, -8, -3, -1.5\}$ (black, maroon, green, red, respectively) in linear (Fig. 1a) and semilogarithmic (Fig. 1b) scales.

The difference between these densities becomes more apparent in semilogarithmic scale, where the limiting distribution (for $\alpha \to -\infty$) appears as a straight line. The larger the α is, the more prone the random variable to produce extreme values is.

Given the sample $\boldsymbol{z} = (z_1, \ldots, z_n)$ of independent and identically distributed random variables with common distribution $\mathcal{G}^0(\alpha, \gamma, L)$ with $(\alpha, \gamma) \in \Theta, \Theta = \mathbb{R}_- \times \mathbb{R}_+$, a maximum likelihood estimator of (α, γ) satisfies

$$(\widehat{\alpha}, \widehat{\gamma}) = \arg \max_{(\alpha,\gamma)\in\Theta} \mathcal{L}(\alpha, \gamma, L, \boldsymbol{z}),$$

where $\mathcal{L}$ is the likelihood function under the $\mathcal{G}^0(\alpha, \gamma, L)$ distribution. This leads to $\widehat{\alpha}$ and $\widehat{\gamma}$ such that

$$n[\Psi^0(-\widehat{\alpha}) - \Psi^0(L - \widehat{\alpha})] + \sum_{i=1}^{n} \ln \frac{\widehat{\gamma} + L z_i^2}{\widehat{\gamma}} = 0 \tag{3}$$

$$\frac{n\widehat{\alpha}}{\widehat{\gamma}} + (L - \widehat{\alpha}) \sum_{i=1}^{n} (\widehat{\gamma} + Lz_i)^{-1} = 0, \tag{4}$$

where $\Psi^0(t) = \mathrm{d} \ln \Gamma(t)/\mathrm{d}t$ is the digamma function. In many cases, no explicit solution for this system is available and numerical methods have to be used. In this work, we applied the BFGS (Broyden 1965) optimization algorithm.

3 Geodesic distance between $\mathcal{G}^0$ models

Naranjo-Torres et al. (2017) obtained two cases of geodesic distances between $\mathcal{G}^0$ distributions with a known number of looks: the cases where either the texture or the scale is known. These are given, respectively, by

$$s(\alpha_1, \alpha_2) = \left| \int_{\alpha_1}^{\alpha_2} \sqrt{\sum_{n=1}^{L} (-\alpha + n - 1)^{-2}} \, \mathrm{d}\alpha \right|, \text{ and by} \tag{5}$$

$$s(\gamma_1, \gamma_2) = \left| \sqrt{\frac{-\alpha L}{-\alpha + L + 1}} \ln \frac{\gamma_1}{\gamma_2} \right|. \tag{6}$$

The first equation can be solved explicitly for $L = \{1, 2\}$:

$$s(\alpha_1, \alpha_2)\Big|_{L=1} = \left| \ln \frac{\alpha_1}{\alpha_2} \right|, \text{ and}$$

$$s(\alpha_1, \alpha_2)\Big|_{L=2} = \left| \ln \frac{\alpha_1^2(\alpha_2 - 1)^2(\alpha_2 R_2 - 1)((\alpha_1 - 1)R_1 + 1)}{\alpha_2^2(\alpha_1 - 1)^2(\alpha_1 R_1 - 1)((\alpha_2 - 1)R_2 + 1)} + \sqrt{2} \ln \frac{1 + \alpha_2(R_2 - 2) - \alpha_2^2 R_2}{1 + \alpha_1(R_1 - 2) - \alpha_1^2 R_1} \right|,$$

where $R_1 = R(\alpha_1)$ and $R_2 = R(\alpha_2)$ are given by

$$R(\alpha) = \sqrt{\frac{4\alpha^2 - 4\alpha + 2}{(\alpha - 1)^2 \alpha^2}}.$$

Notice that $s(\gamma_1, \gamma_2)$ depends on the texture α, while $s(\alpha_1, \alpha_2)$ is independent of the scale γ. Both (5) and (6) depend on the number of looks L.

To the best of the authors' knowledge, there is no closed expression for the geodesic distance between two $\mathcal{G}^0$ models with both α and γ different, given L known.

Both distances can be turned into test statistics (see Verdoolaege and Scheunders 2011; Ilea et al. 2015) by indexing with maximum likelihood estimators based on samples of sizes m and n and then rescaling:

$$T = \frac{mn}{m+n}\widehat{s}^2.$$

We will denote

$$T_\alpha = \frac{mn}{m+n}\ s(\widehat{\alpha}_1, \widehat{\alpha}_2)\big)^2, \text{ and} \tag{7}$$

$$T_\gamma = \frac{mn}{m+n}\ s(\widehat{\gamma}_1, \widehat{\gamma}_2)\big)^2. \tag{8}$$

Under the null hypothesis of equal parameters, when $m, n \to \infty$ proportionally, both T_α and T_γ follow a χ^2_1 distribution, so it is possible to compute the p value of two samples under H_0 and either reject or not this hypothesis (Menendez et al. 1995).

Section 4 presents an analysis of the behavior of these test statistics T_α and T_γ. Section 5 studies ways of combining them to produce a two-parameter test.

4 Analysis of one-parameter tests

In this section, we analyze the finite sample size behavior of the test statistics defined in (7) and (8) using Monte Carlo experiments. We obtained the samples following the guidelines presented in Chan et al. (2018).

The parameter space for the first experiment was $\alpha = -1.5$ and the same sample size $n \in \{50, 100, 150, \ldots, 1000\}$ for $\gamma = 1$ and $L = 1$. We obtained five thousand independent replications for each sample size and maximized the following reduced log-likelihood function:

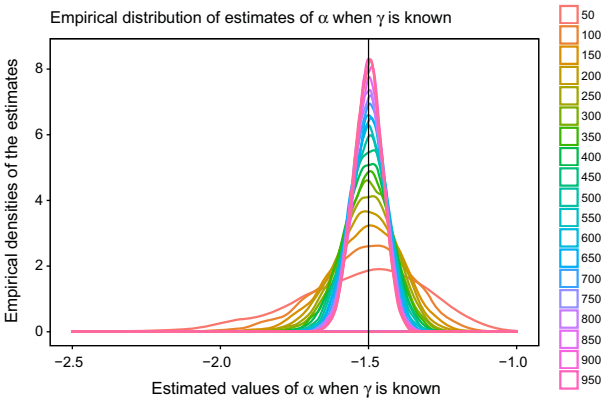
$$\ell(\alpha; \gamma, L, z) = n[\log\Gamma(L-\alpha) - \alpha\log\gamma - \log\Gamma(-\alpha)] + \alpha\sum_{i=1}^{n}\log(\gamma + Lz_i). \tag{9}$$

We produced two independent samples in each replication in order to compute a distance from the respective estimated models.

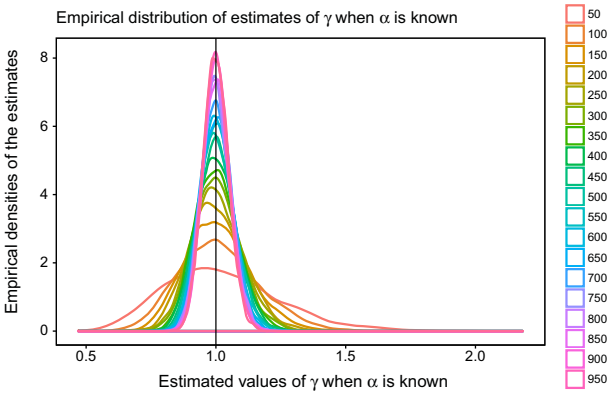
Figure 2a presents the sample densities of $\widehat{\alpha}$ for $\gamma = 1$ and $L = 1$. They are all centered around the true value $\alpha = -1.5$, and as expected, the larger the sample size n is, the smaller the variability is. Small values of n yield more asymmetric densities than their larger counterpart. The parameter space and number of replications for the second experiment were the same, but the reduced log-likelihood to be maximized was

$$\ell(\gamma; \alpha, L, z) = -n\alpha\log\gamma + (\alpha - L)\sum_{i=1}^{n}\log(\gamma + Lz_i). \tag{10}$$

Similar conclusions can be drawn from the sample densities of the maximum likelihood estimators of γ, when α and L are known; cf. Fig. 2b.



(a) Sample densities of $\widehat{\alpha}$.

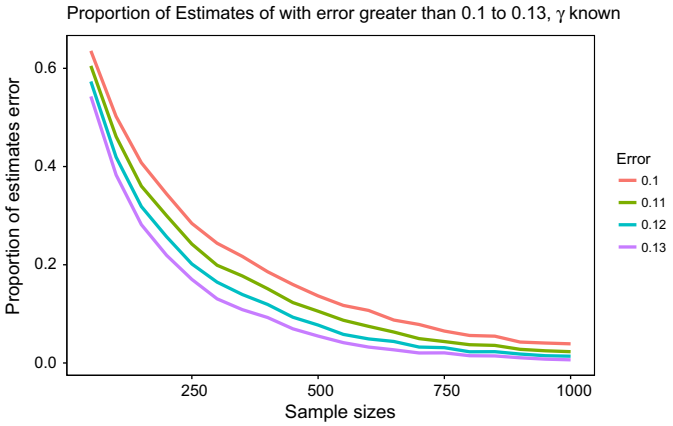


(b) Sample densities of $\widehat{\gamma}$.

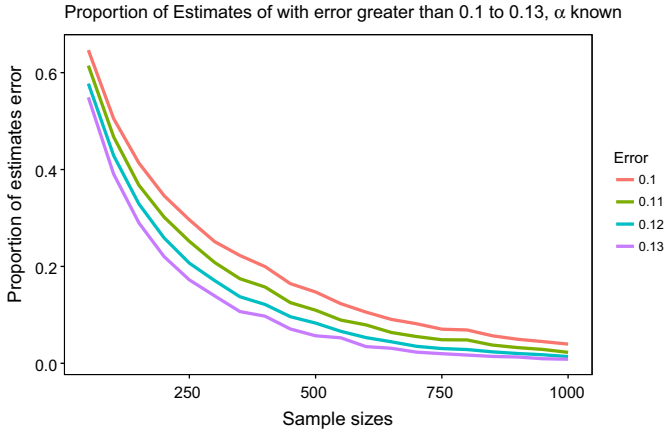
Fig. 2 Estimated densities of maximum likelihood estimators of α and γ, when only one parameter is unknown and $L = 1$

The behavior shown in Fig. 2 is consistent across other values of α and γ. Both (9) and (10), as well as the two-parameter reduced log-likelihood function presented below, were optimized using the `maxLik` routine (Henningsen and Toomet 2011) available in `R` (R Core Team 2016).

Figure 3a, b shows the proportion of estimates whose error is larger than 0.10, 0.11, 0.12, 0.13, when the γ parameter is known and when the α parameter is known, respectively. The experiment consists of generating 5000 samples of size $n = \{50, 100, 150, \ldots, 1000\}$, with $\mathcal{G}_I^0(\alpha, \gamma, L)$ distribution. In this case, $\alpha = -1.5$, $\gamma = 1$, $L = 1$. It can be observed that the proportion of estimates with error dramatically decreases as the sample size increases. This evidences that bias of the maximum likelihood estimates strongly depends on the sample size. Sample sizes greater than



(a) Proportion of estimates of α whose error is larger than 0.10, 0.11, 0.12, 0.13.



(b) Proportion of estimates of γ whose error is larger than 0.10, 0.11, 0.12, 0.13.

Fig. 3 Proportion of α and γ test statistics with errors larger than 0.10, 0.11, 0.12, 0.13

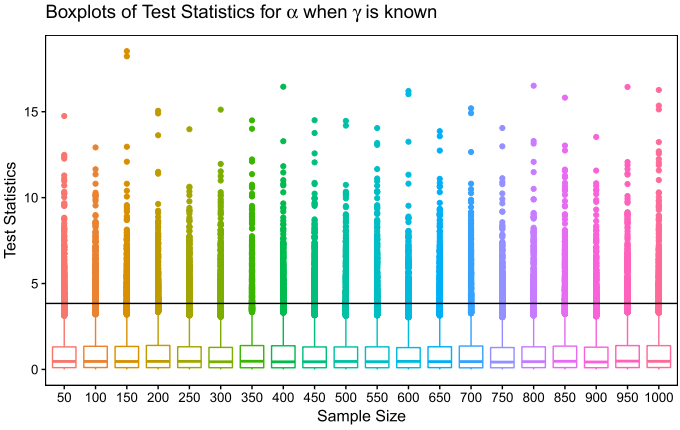
or equal to 750 provide acceptable results, but in practical situations, one is often interested in smaller samples, e.g., for filters which compute estimates over windows of size 7×7. The selected values of the parameters are arbitrary, in order to show an example of the maximum likelihood estimator behavior as the sample size increases.

As said, in each replication two independent samples were generated, and an estimate computed with each. Each pair of estimates is then used to compute either T_α or T_γ, depending on the experiment. Our main interest lies in the finite sample behavior of these test statistics.

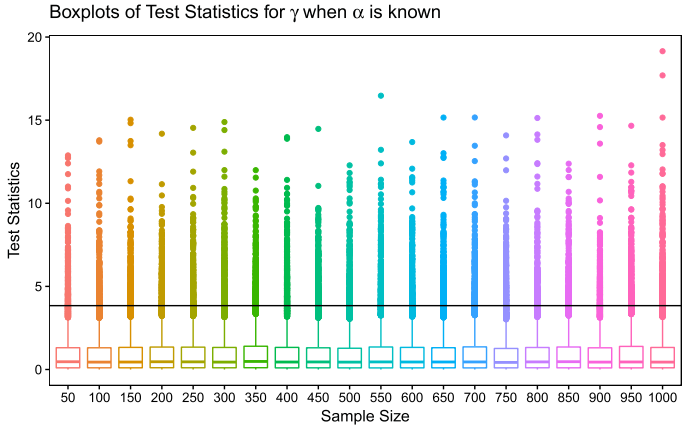
4.1 Finite sample size behavior of T_α

For each sample size, we have five thousand samples of T_α. We will analyze the distribution of these test statistics and the empirical size of the test when compared with the asymptotic result.

Figure 4a shows the boxplots of the T_α test statistics for different sample sizes, along with the theoretical cut value at 95 % (approximately 3.841 459, the 0.95 quantile of the χ^2_1 distribution).



(a) Boxplots of T_α for γ known.



(b) Boxplots of T_γ for α known.

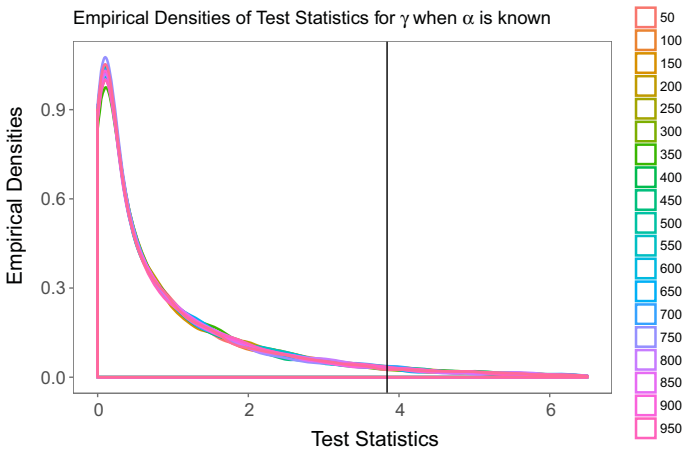
Fig. 4 Boxplots of tests statistics

Figure 5a shows the sample densities of the T_α test statistics for different sample sizes, along with the theoretical cut value at the 95 % (approximately 3.841 459, the 0.95 quantile of the χ^2_1 distribution).

Neither Fig. 4a nor Fig. 5a suggests any significant change of distribution of T_α when the sample size varies, an evidence that $n = 50$ is a large enough sample size to attain the asymptotic properties.

Figure 6a presents the empirical size of T_α tests for different sample sizes, along with the theoretical cut value. The minimal and maximal deviation between the empirical and theoretical p values is, respectively, 0.4 % and 13.2 %.

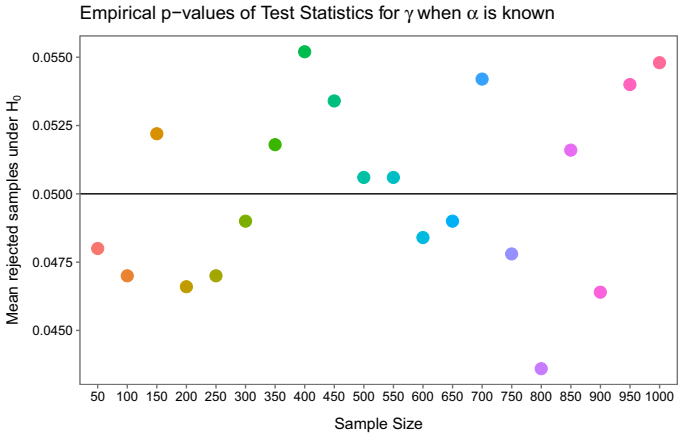
(a) Empirical densities of T_α, for γ known.



(b) Empirical densities of T_γ, for α known.

Fig. 5 Empirical densities of tests statistics

(a) Empirical p-values of T_α, γ known.



(b) Empirical p-values of T_γ, α known.

Fig. 6 Empirical size of tests statistics

4.2 Finite sample size behavior of T_γ

For each sample size, we have five thousand samples of T_γ. We will analyze the distribution of these test statistics and the empirical size of the test with respect to the asymptotic value.

Figure 4b shows the boxplots of the T_γ test statistics for different sample sizes, along with the theoretical cut value at the 95 % (approximately 3.841 459, the 0.95 quantile of the χ^2_1 distribution). Figure 5b shows the sample densities of the T_γ test statistic, for different sample sizes, along with the theoretical cut value at the 95 % (approximately 3.841 459, the 0.95 quantile of the χ^2_1 distribution).

Neither Fig. 4b nor Fig. 5b suggests any significant change of distribution of T_γ when the sample size varies, evidence that $n = 50$ is a large enough sample size to attain the asymptotic properties. This motivates the use of a single model, namely the χ^2_1 distribution, for computing quantiles.

Figure 6b presents the empirical size of T_γ test for different sample sizes, along with the theoretical cut value. The minimal and maximal deviation between the empirical and theoretical p values is, respectively, 1.2 % and 12.8 %.

5 Analysis of two-parameter tests

In this section, we analyze the more realistic situation of estimating both the scale and texture parameters, while assuming the number of looks known. As mentioned, we opted for computing $(\widehat{\alpha}, \widehat{\gamma})$ the maximum likelihood estimator of (α, γ) by maximizing the reduced log-likelihood function which, for L known, is

$$\begin{aligned}\ell(\alpha, \gamma; L, \boldsymbol{z}) = n[&\log\Gamma(L-\alpha) - \alpha\log\gamma - \log\Gamma(-\alpha)] \\ &+ (\alpha - L)\sum_{i=1}^{n}\log(\gamma + Lz_i). \end{aligned} \tag{11}$$

Again, the routine `maxLik` was the tool employed for maximizing (11).

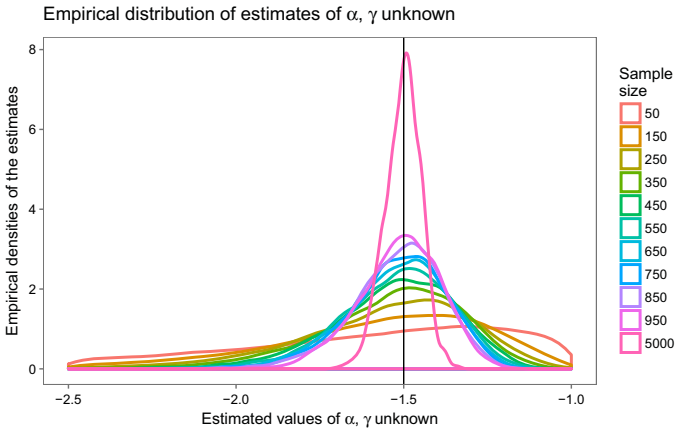
Whereas maximizing (9) and (10) poses no numerical problem, (11) has well-reported problems caused by cases where this likelihood becomes flat (Frery et al. 2004). In order to avoid such problems without introducing specialized techniques that depart from the concept of maximum likelihood, only solutions satisfying $(\widehat{\alpha}, \widehat{\gamma}) \in [15\alpha, 0) \times (0, 15\gamma]$ were considered feasible. The number of replications is computed over feasible solutions.

The parameter space of the study is the product of the sets $\alpha \in \{-1.5, -3, -4\}$, $L \in \{1, 2\}$, and $n \in \{50(100)950, 5000\}$. For each α, the scale is $\gamma = -\alpha - 1$, so the expected value is 1. Following the recommendations discussed in Bustos and Frery (1992), the number of replications changes with the sample size as $R = [R_{\max}/n]$; we empirically found $R_{\max} = 5 \times 10^6$ produces reliable results with an acceptable computational cost.

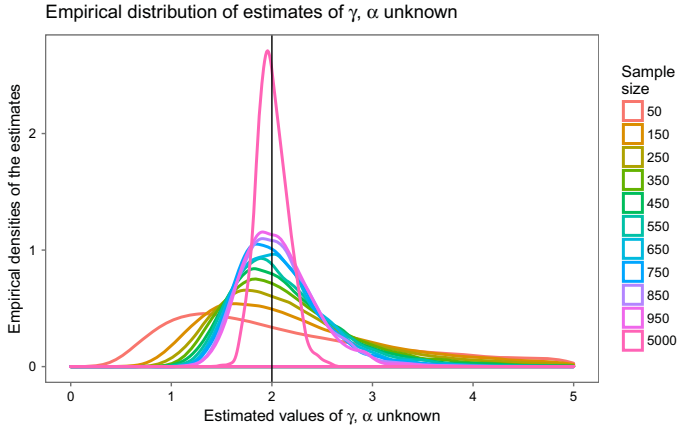
The plots in Fig. 7 show the empirical densities of the estimators of texture and scale, Fig. 7a for the case $\alpha = -1.5$, $\gamma = 0.5$ and $L = 1$, Fig. 7b for the case $\alpha = -3$, $\gamma = 2$ and $L = 1$.

The difference between Figs. 2a and 7a is noticeable in terms of spread and centrality. The same observation holds when comparing Figs. 2b and 7b. The effect of missing the information of one parameter is, thus, remarkable.

Figure 8 shows the contour plots of the estimates $(\widehat{\alpha}, \widehat{\gamma})$ for samples of size $n = 50$ and all the cases here considered. This figure corroborates that it is not adequate to assume that $\widehat{\alpha}$ and $\widehat{\gamma}$ can be uncorrelated, let alone independent.



(a) Empirical densities of $\widehat{\alpha}$.



(b) Empirical densities of $\widehat{\gamma}$.

Fig. 7 Empirical densities of estimators when only $L = 1$ is known, $\alpha = -3$ and $\gamma = 2$

This relationship between estimators is also exhibited by the tests statistics that use them. Figure 9 shows the contour plots of (T_α, T_γ).

In practice, one needs to discriminate regions with unknown texture and scale, so a test statistic for both parameters, say $T_{\alpha,\gamma}$, is required.

The strong relationship between $\widehat{\alpha}$ and $\widehat{\gamma}$ is evident, so is the same relationship between test statistics, therefore just adding T_α and T_γ and assuming that the sum follows a χ^2_2 law might not be a good idea. This justifies the following analysis which aims at finding relevant properties of two-to-one transformations of $(T_\alpha, T_\gamma) \to T_{\alpha,\gamma}$, in search for a test statistic for assessing the null hypothesis of having two samples from the same $\mathcal{G}^0$ distribution. We analyze the following test statistics:

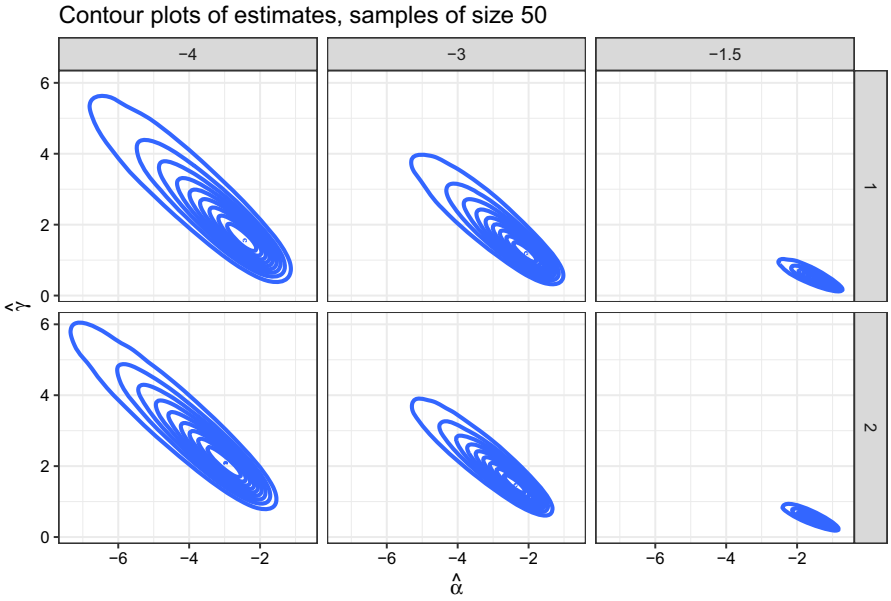


Fig. 8 Contour plots of $(\widehat{\alpha}, \widehat{\gamma})$ for all the cases considered and samples of size $n = 50$

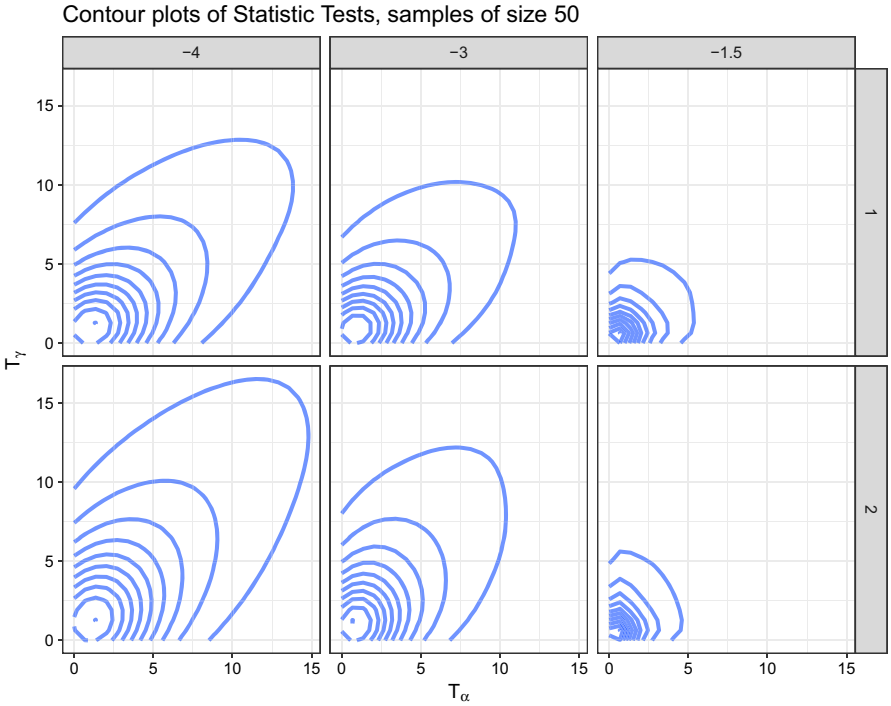


Fig. 9 Contour plots of (T_α, T_γ) for all the cases considered and samples of size $n = 50$

$$T^1_{\alpha,\gamma} = \sqrt{T^2_{\alpha=(\widehat{\alpha}_1+\widehat{\alpha}_2)/2} + T^2_{\gamma}}, \tag{12}$$

$$T^2_{\alpha,\gamma} = \frac{T_{\alpha=(\widehat{\alpha}_1+\widehat{\alpha}_2)/2} + T_{\gamma}}{2}, \tag{13}$$

$$T^3_{\alpha,\gamma} = \max\left\{\frac{T_{\alpha=(\widehat{\alpha}_1+\widehat{\alpha}_2)/2}}{T_{\gamma}}, \frac{T_{\gamma}}{T_{\alpha=(\widehat{\alpha}_1+\widehat{\alpha}_2)/2}}\right\} \tag{14}$$

Equations (12), (13) and (14) are combinations of statistics for a single free parameter, given in Eqs. (7) and (8), but their distributions are unknown, and thus, we cannot apply a hypothesis test to decide if two samples come from the same distribution or not. So, to solve this problem, we estimate these distributions using permutation methods, as explained in Sect. 5.1.

5.1 Permutation methods

Permutation methods are a type of statistical significance test which can be applied to statistics with unknown distribution. They were developed by Fisher (1934). Berry et al. (2011, 2018) explain the advantages of this type of tests. There are at least two kinds of permutation tests:

Exact: in which all possible reorganizations of the sample are considered. This kind has high computational cost, depending on the sample size.

Random: which consider a certain amount of permutations, usually 1000 or 10,000. They are more appropriate if the sample size is large.

In this work, we test if two samples $X \sim \mathcal{G}^0(\alpha_1, \gamma_1, L)$ and $Y \sim \mathcal{G}^0(\alpha_2, \gamma_2, L)$ are from the same distribution, then we pose the null hypothesis H_0:$(\alpha_1, \gamma_1) = (\alpha_2, \gamma_2)$ and we want to know the probability of rejecting it. With this objective, we estimate the empirical distributions of the tests $T^i_{\alpha,\gamma}$, $i = 1, 2, 3$ from Eqs. (12), (13) and (14) by means of the following steps. For more information, see Feinstein (1993).

1. Choose a statistic $T^i_{\alpha,\gamma}$, $i = 1, 2, 3$ from Eqs. (12), (13) or (14).
2. Generate z_1 and z_2 random samples of sizes m and n, respectively, both from the same $\mathcal{G}^0(\alpha, \gamma, L)$ distribution granting the null hypothesis. Let perm be the number of permutations; in our experiment perm $= 1000$.
3. Compute the estimates $(\widehat{\alpha}_1, \widehat{\gamma}_1)$ and $(\widehat{\alpha}_2, \widehat{\gamma}_2)$ with each sample.
4. Calculate the observed statistic value, $T^i_{\alpha,\gamma}$, with the data from z_1 and z_2.
5. Repeat for $k = 1, \ldots,$ perm:
   - Shuffle the joint sample $z = (z_1, z_2)$ and divide it in two groups of sizes m and n, say z^k_1 and z^k_2, respectively.
   - Compute the estimates $(\widehat{\alpha}_1, \widehat{\gamma}_1)$ and $(\widehat{\alpha}_2, \widehat{\gamma}_2)$ for each sample z^k_1 and z^k_2.
   - Calculate the statistic value using the permuted samples, $T^p_{\alpha,\gamma}(k)$.
   - Compare the observed statistic value calculated in Step 4 with the statistic computed after permutation $T^p_{\alpha,\gamma}(k)$.

6. The proportion of differences equal to or larger than the observed statistic value serves as the p value for the permutation test, or:

$$p \text{ value} = \frac{\#\{k : T^{p}_{\alpha,\gamma}(k) \geq T^{i}_{\alpha,\gamma}\}}{\text{perm}}. \tag{15}$$

7. If p value $< \eta$, the null hypothesis is rejected at level η.

5.2 Results of applying permutation methods

We use devised Monte Carlo experiments to quantify empirical rejection rate (R-rate) generated by the proposed tests, under the null hypothesis. The experiment is repeated 500 times.

Table 1 shows the results of applying the permutation test to the statistic given in Eqs. (12), (13) and (14), for values of $L = \{1, 2\}$, $\alpha = \{-1.5, -4\}$ and $\gamma = -\alpha - 1$, at level $\eta = 0.05$. For lack of space, we present only the results for $n = 50$, $n = 550$ and $n = 5000$, corresponding to small, medium and large samples. We inform the rejection rate under the null hypothesis (false negative rate) which is the estimated test size. Tests T^1 and T^2 exhibit the closest empirical sizes to the nominal level. It can be observed that if the sample size increases, the false negative rate is not necessarily reduced.

Figure 10 shows the false negative rate for the test in Eq. 14, under the null hypothesis depending on the sample size, for $\alpha = -1.5$, $\gamma = 0.5$, $L = 1$. It can be observed that the false negative rate fluctuates around the value of the level $\eta = 0.05$ represented with a green straight line and the highest value of the false negative rate is given for the sample size $n = 50$.

Table 1 Rejection rates for the proposed statistics under the null hypothesis

L	α	n	R-rate T^1	R-rate T^2	R-rate T^3
1	−1.5	50	0.048	0.058	0.075
		550	0.056	0.050	0.051
		5000	0.044	0.048	0.058
1	−4	50	0.046	0.046	0.049
		550	0.056	0.056	0.045
		5000	0.046	0.046	0.050
2	−1.5	50	0.060	0.056	0.05
		550	0.052	0.052	0.043
		5000	0.052	0.042	0.059
2	−4	50	0.06	0.060	0.051
		550	0.038	0.038	0.035
		5000	0.048	0.048	0.055

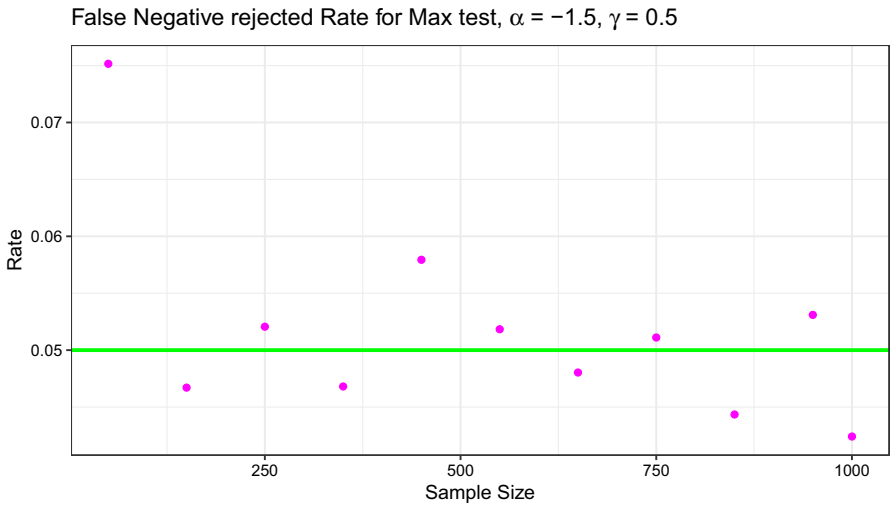


Fig. 10 False negative rate depending on the sample size, under the null hypothesis

5.3 Application in edge detection

In this section, we present an application of the proposed method to the problem of edge detection in actual SAR images. Gambini et al. (2008) proposed a general and flexible algorithm for edge detection which is based on finding, in a narrow strip of data, the point where there is maximum evidence of a change of properties. Naranjo-Torres et al. (2017) used a geodesic distance between models as a measure of this change, assuming the $\mathcal{G}_I^0$ distribution with known scale parameter. In this work, we use the same algorithm but considering two parameters unknown: texture α and scale γ. In order that this work is self-contained, we briefly explain the algorithm. For more information, see Naranjo-Torres et al. (2017).

Let I be an actual SAR image of m lines and n columns of pixels. In this application, we use only one line of data, i.e., a strip of size $1 \times n$. In each step $3 \leq k \leq n-3$, we divide the line in two disjoint samples, $S_1(k) = (z_1, \ldots, z_k)$ and $S_2(k) = (z_{k+1}, \ldots, z_n)$ used to estimate the parameters $(\widehat{\alpha}_1, \widehat{\gamma}_1)(k)$ and $(\widehat{\alpha}_2, \widehat{\gamma}_2)(k)$, respectively, by maximum likelihood. Then, the p value $p(k)$ is computed using the method described in Sect. 5.1.

Finally, we estimate the transition point as the position at which $p(k)$ is minimum: $\widehat{col} = \arg\min_k p(k)$. The method is sketched in Algorithm 1, where I is the original image and m and n are the numbers of rows and columns of the input image. Notice that the minimum sample size is set to three observations.

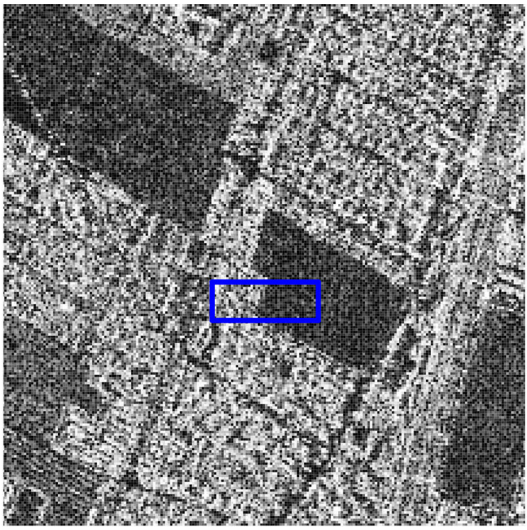
Figure 11 shows the results of applying the edge detector algorithm. Figure 11a shows the SAR image and presents the area where the edge detection was performed. Figure 11b shows the result of applying the edge detector to each line in a selected region.

Algorithm 1: Edge Detection by the geodesic distance of the $\mathcal{G}_I^0$ distribution with two unknown parameters.

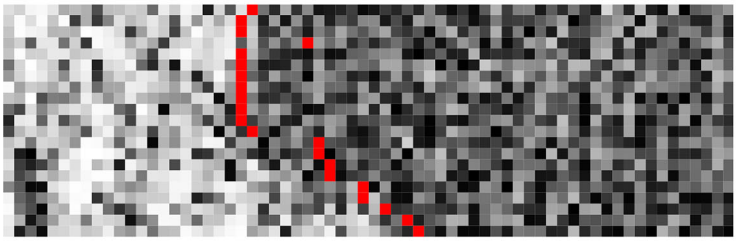
1: input: I, m, n
2: **for** each line of I, $i = 1, \ldots, m$ **do**
3: **for** $k = 3, \ldots, n-3$ **do**
4: Divide the line in two samples $S_1(k) = (z_1, \ldots, z_k)$ and $S_2(k) = (z_{k+1}, \ldots, z_n)$.
5: Estimate (α, γ) by maximum likelihood in each sample, obtaining $(\widehat{\alpha}_1, \widehat{\gamma}_1)(k)$ and $(\widehat{\alpha}_2, \widehat{\gamma}_2)(k)$.
6: Compute $T(k) = T_{\widehat{\alpha}(k), \widehat{\gamma}(k)}$ using Eqs. (12), (13) or (14).
7: Consider the array of statistics between the pairs of samples: $\boldsymbol{T} = \{T(k),\ 3 \leq k \leq m-3\}$ and compute the array of p values $\boldsymbol{P} = \{p(k),\ 3 \leq k \leq m-3\}$.
8: Find the column where the array $\boldsymbol{P}$ is minimized, which corresponds to the transition point on the line i:

$$\widehat{col} = \arg\min_k \boldsymbol{p}(k),$$

9: **end for**
10: **end for**



(a) SAR image and the region used.



(b) Edge points found over each line of the region.

Fig. 11 Results of applying the edge detector to actual data using $T^1_{\alpha,\gamma}$

6 Conclusions and future work

Unable to calculate the geodesic distance of the $\mathcal{G}^0$ distribution depending on two free parameters, we carried out a study dedicated to evaluating the possibility of using a combination of tests based on the geodesic distance with a single unknown parameter, as calculated in Naranjo-Torres et al. (2017).

We compare three statistics whose distributions are unknown. We use permutation methods to estimate their empirical distributions.

The results show that, under the null hypothesis, the false negative rate fluctuates around the rejection level, even with small samples. It can be observed that if the sample size increases, the false negative rate is not necessarily reduced; this encourages us to continue the investigations with small samples. The results are promising and can be readily employed in speckled image processing and analysis.

Acknowledgements Funding was provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, BR) and Fundação de Amparo à Pesquisa do Estado de Alagoas (Fapeal, BR).

Appendix

Simulations were performed using the `R` language and environment for statistical computing version 3.0.2 (R Core Team 2016).

References

Atkinson C, Mitchell AF (1981) Rao's distance measure. Sankhyā: Indian J Stat Ser A (1961-2002) 43:345–365

Berry KJ, Johnston JE, Mielke PW (2011) Permutation methods. Wiley Interdiscip Rev Comput Stat 3(6):527–42. https://doi.org/10.1002/wics.177

Berry KJ, Johnston JE, Mielke PW, Johnston LA (2018) Permutation methods. Part II. Wiley Interdiscip Rev Comput Stat 10:e1429. https://doi.org/10.1002/wics.1429

Broyden CG (1965) A class of methods for solving nonlinear simultaneous equations. Math Comput 19:577–593

Bustos OH, Frery AC (1992) Reporting Monte Carlo results in statistics: suggestions and an example. Rev Soc Chil Estad 9(2):46–95

Chan D, Rey A, Gambini J, Frery AC (2018) Sampling from the GI0 distribution. Monte Carlo Methods Appl 24(4):271–287

Dell'Acqua F, Gamba P (2012) Remote sensing and earthquake damage assessment: experiences, limits, and perspectives. Proc IEEE 100(10):2876–2890

Feinstein AR (1993) Permutation tests and statistical significance. M. D. Comput Comput Med Pract 10:28–41

Fisher KA (1934) Statistical methods for research workers, 5th edn. Oliver & Boyd, Edinburgh

Frery AC, Müller H-J, Yanasse CCF, Sant'Anna SJS (1997) A model for extremely heterogeneous clutter. IEEE Trans Geosci Remote Sens 35(3):648–659

Frery AC, Cribari-Neto F, Souza MO (2004) Analysis of minute features in speckled imagery with maximum likelihood estimation. EURASIP J Appl Signal Process 2004(16):2476–2491

Gambini J, Mejail M, Jacobo-Berlles J, Frery A (2006) Feature extraction in speckled imagery using dynamic B-spline deformable contours under the $\mathcal{G}^0$ model. Int J Remote Sens 27(22):5037–5059

Gambini J, Mejail M, Jacobo-Berlles J, Frery AC (2008) Accuracy of edge detection methods with local information in speckled imagery. Stat Comput 18(1):15–26

Gambini J, Cassetti J, Lucini M, Frery A (2015) Parameter estimation in SAR imagery using stochastic distances and asymmetric kernels. IEEE J Sel Top Appl Earth Obs Remote Sens 8(1):365–375
Henningsen A, Toomet O (2011) maxLik: a package for maximum likelihood estimation in R. Comput Stat 26(3):443–458
Hill MJ, Ticehurst CJ, Lee J-S, Grunes MR, Donald GE, Henry D (2005) Integration of optical and radar classifications for mapping pasture type in Western Australia. IEEE Trans Geosci Remote Sens 43(7):1665–1681
Ilea I, Bombrun L, Germain C, Terebes R, Borda M (2015) Statistical hypothesis test for robust classification on the space of covariance matrices. In: IEEE international conference on image processing (ICIP), pp 271–275
Marques RCP, Medeiros FN, Santos Nobre J (2012) SAR image segmentation based on level set approach and $\mathcal{G}_A^0$ model. IEEE Trans Pattern Anal Mach Intell 34(10):2046–2057
Menendez ML, Morales D, Pardo L, Salicru M (1995) Statistical test based on the geodesic distances. Appl Math Lett 8(1):65–69
Naranjo-Torres J, Gambini J, Frery AC (2017) The geodesic distance between $\mathcal{G}_I^0$ models and its application to region discrimination. IEEE J Sel Top Appl Earth Obs Remote Sens 10(3):987–997
Nascimento ADC, Cintra RJ, Frery AC (2010) Hypothesis testing in speckled data with stochastic distances. IEEE Trans Geosci Remote Sens 48(1):373–385
Quartulli M, Datcu M (2004) Stochastic geometrical modelling for built-up area understanding from a single SAR intensity image with meter resolution. IEEE Trans Geosci Remote Sens 42(9):1996–2003
R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
Rao CR (1945) Information and the accuracy attainable in the estimation of statistical parameters. Bull Calcutta Math Soc 37:81–91
Rao CR (1992) Information and the accuracy attainable in the estimation of statistical parameters. In: Kotz S, Johnson NL (eds) Breakthroughs in statistics. Springer series in statistics. Springer, New York, pp 235–247
Salicrú M, Morales D, Menéndez ML, Pardo L (1994) On the applications of divergence type measures in testing statistical hypotheses. J Multivar Anal 51(2):372–391
Silva WB, Freitas CC, Sant'Anna SJS, Frery AC (2013) Classification of segments in PolSAR imagery by minimum stochastic distances between Wishart distributions. IEEE J Sel Top Appl Earth Obs Remote Sens 6(3):1263–1273
Storie CD, Storie J, Salinas de Salmuni G (2012) Urban boundary extraction using 2-component polarimetric SAR decomposition. In: IEEE international geoscience and remote sensing symposium (IGARSS), pp 5741–5744
Sun W, Shi L, Yang J, Li P (2016) Building collapse assessment in urban areas using texture information from postevent SAR data. IEEE J Sel Top Appl Earth Obs Remote Sens 9(8):3792–3808
Verdoolaege G, Scheunders P (2011) Geodesics on the manifold of multivariate generalized Gaussian distributions with an application to multicomponent texture discrimination. Int J Comput Vis 95(3):265–286
Verdoolaege G, Scheunders P (2012) On the geometry of multivariate generalized Gaussian models. J Math Imaging Vis 43(3):180–193