Analysis of Speckled Imagery with Parametric and Nonparametric Tests

Abraão D. C. Nascimento  
Department of Statistics  
Federal University of Pernambuco  
Recife, Pernambuco, Brazil  
abraao.susej@gmail.com

Alejandro C. Frery  
Department of Statistics  
Federal University of Pernambuco  
Recife, Pernambuco, Brazil  
acfrery@pesquisador.cnpt.br

Renato J. Cintra  
Department of Statistics  
University of Calgary  
Calgary, Alberta, Canada T2N 1N4  
rjds@de.ufpe.br

Abstract—Synthetic aperture radar (SAR) has a pivotal role as a remote imaging method. Obtained by means of coherent illumination, SAR images are contaminated with speckle noise. The statistical modelling of such contamination is well described according the multiplicative model and its implied $\mathcal{G}^0$ distribution. The understanding of SAR imagery and scene element identification is an important objective in the field. In particular, reliable image contrast tools are sought. Aiming the proposition of new tools for evaluating SAR image contrast, we investigated method based on stochastic divergence based. As a consequence, we proposed several divergence measures specifically tailored for $\mathcal{G}^0$ distributed data. We also introduced a nonparametric approach based on the Kolmogorov-Smirnov distance for $\mathcal{G}^0$ data. Statistic tests based on such measures were also devised and assessed. Their performance were quantified according to the resulting test sizes and powers. Using a Monte Carlo simulation approach, a robustness analysis is presented for several degrees of contamination. It was identified that the proposed tests based on the triangular and the arithmetic-geometric measures outperformed the classical Kolmogorov-Smirnov, Kullback-Leibler and Bhattacharyya approaches.

Index Terms—distance; robust; parametric; nonparametric; multiplicative model;

I. INTRODUCTION

Synthetic aperture radar (SAR) images are obtained with coherent illumination and, consequently, their appearance is affected by a signal-dependent granular noise called speckle [1]. Due to deviations from classical properties of additivity and Gaussian distribution, speckled images require tailored processing and analysis techniques which could conveniently rely on the statistical properties of the data.

The ability to discriminate regions in an image is an important tool for identifying and classifying distinct targets in the scene. This fact has motivated the proposal of statistically based contrast measures [2]. Therefore, an accurate modelling of speckled data is a pivotal preprocessing step for SAR image analysis [3]. Indeed, the probability distribution of the image data is a crucial information for any subsequent statistical analysis.

Among the proposals for SAR image modelling, the multiplicative model [3] has proven to be a important framework for statistically processing and analyzing SAR data [4]. Assuming that the hypothesis suggested by the multiplicative model are valid, the $\mathcal{G}^0$ distribution arises as an efficient and effective probabilistic model for speckled data [5], [6].

Hypothesis test methods have been considered as a venue to quantifying the image contrast between different SAR regions. Edge detection [7], [8] and classification procedures [9] rely on parametric and nonparametric hypothesis tests. Although some nonparametric methods are computationally convenient, they seldom take explicitly the roughness information into account [7]. As a consequence, simple measures are unable to provide any information about that important parameter. Additionally, these methods cannot offer any known statistical property that could be utilized for the design of hypothesis testing procedures [8]. Thus, the proposal of hypothesis testing methods based on roughness dependent contrast measures are a much sought tool for speckled data analysis.

In recent years, the interest in adapting information-theoretic tools to image processing has increased notably. In [2], the Bhattacharyya distance was applied as a scalar contrast measure for polarimetric and interferometric SAR imagery. In a previous work [10], [11], several parametric methods based on $(h, \phi)$-divergence measures were proposed and submitted to a comprehensive examination.

Often regarded to as a classical benchmark, the Kolmogorov-Smirnov distance has been used for segmentation and labelling of SAR images [12]. This nonparametric test has found applications in polarimetric speckled data analysis, for instance [9]. Additionally, the Kolmogorov-Smirnov distance was also applied as a goodness-of-fit measure for SAR data modelling [13].

Whereas on the one hand a careful statistical modeling is fundamental, on the other hand nature is more complex than most models. Even sophisticated statistical models and tools may have their effectiveness diminished when data deviate from the assumed underlying hypothesis, even mildly. Robust statistics is a way of tackling this problem, and resistant techniques have been used with success in the SAR literature [14]–[17].

The aim of this work is to assess the SAR image con-
tract capabilities of selected parametric methods based on divergences measures, when compared to the nonparametric Kolmogorov-Smirnov testing methodology. Such capability will be described by means of the size and the power of hypothesis tests, under several situations which include “pure” and contaminated data. An analytical expression for Kolmogorov-Smirnov distance between \( G^0 \) distributions is also presented. Similar results were not found in the literature.

The paper unfolds as following. Section II discusses the multiplicative model and the distribution that will be used therein. Section III presents the distances whose performance will be tested, namely those derived from \((h, \phi)\)-divergences and the Kolmogorov-Smirnov distance. Section IV present the methodology, namely, the Monte Carlo experiments and the results. Section V discusses the conclusions.

II. THE MULTIPLICATIVE MODEL AND THE \( G^0 \) DISTRIBUTION

Unlike traditional models, speckle noise has a multiplicative nature and it is strongly non-Gaussian. The multiplicative model is one of the most successful approaches for the description of such noise mainly because of its phenomenological nature which is closely tied to the physics of the image formation [1].

Complex data are usually generated by coherent radiation to form the image. However, the obtained imagery associated to the returning echo is usually described in terms of its amplitude or intensity [3]. In this work, without loss of generality [1], the intensity format is adopted.

Considering positive and independent random variables \( X \) and \( Y \) as models for the terrain backscatter and the speckle noise, respectively, the model for the return is \( Z = X \cdot Y \).

The generalized inverse Gaussian distribution is proposed as a general model for the intensity backscatter, however a tractable particular case is the reciprocal gamma law [3], whose density function is given by

\[
f_X(x; \alpha, \gamma) = \frac{\gamma^{-\alpha} \Gamma(-\alpha)}{\Gamma(-\alpha)} x^{-\alpha-1} \exp \left( -\frac{\gamma}{x} \right), \quad -\alpha, \gamma, x > 0. \tag{1}
\]

When quadratic detection is used, the random variable \( Y \) obeys the gamma distribution [3], with density furnished by

\[
f_Y(y; L) = \frac{L^L y^{L-1} \exp(-Ly)}{\Gamma(L)}, \quad y > 0, L \geq 1, \tag{2}
\]

where \( L \) is the number of looks. Throughout this paper, the number of looks is assumed to be known and constant over the whole image. A detailed account of the, until recently largely unexplored, issue of estimating \( L \) is provided in [4].

Considering independent random variables with distributions characterized by the densities described in Equations (1) and (2), we obtain that the density of \( Z \) is expressed by

\[
f_Z(z; \alpha, \gamma, L) = \frac{L^L \Gamma(-\alpha)}{\Gamma(-\alpha) \Gamma(L)} z^{L-1} (\gamma + Lz)^{-\alpha-L}, \quad -\alpha, \gamma, z > 0, L \geq 1.
\]

We indicate this situation as \( Z \sim G^0(\alpha, \gamma, L) \). As shown in [18], [19], this distribution can be used as a universal model for speckled data.

The \( r \)-th moment of \( Z \) is expressed by

\[
E[Z^r] = \left( \frac{\gamma}{L} \right)^r \frac{\Gamma(-\alpha - r) \Gamma(L + r)}{\Gamma(-\alpha) \Gamma(L)}, \tag{3}
\]

if \(-r/2 > \alpha\). Otherwise, it is infinite.

Figure 1 shows nine \( 256 \times 256 \) pixel regions of simulated speckled data according to the \( G^0 \) distribution for \( L = 4 \), \( \alpha \in \{-8, -5, -3\} \), and \( \gamma \in \{6, 12, 120\} \). The homogeneity of the resulting regions depends on the choice of \( \alpha \), the roughness parameter, while \( \gamma \) controls the mean brightness. Figures 2(a) and 2(b) show the density of \( G^0 \) laws with varying \( \alpha \in \{-3, -5, -8, -11\} \) and \( \gamma \in \{6, 12, 30, 60\} \) for the selected values of \( \gamma = 12 \) and \( \alpha = -4 \), respectively. In both cases, the number of looks was considered constant \( L = 4 \). An increase of \( \alpha \) or \( \gamma \) causes the flattening of the associated densities.

Fig. 1. Synthetic \( G^0 \) imagery for \( \gamma \in \{6, 12, 120\} \) (left to right) and \( \alpha \in \{-8, -5, -3\} \) (top to bottom).

Fig. 2. \( G^0 \) densities: (a) \( G^0(\alpha, 12, 4) \), (b) \( G^0(-4, \gamma, 4) \).
Several methods for estimating $\alpha$ and $\gamma$ are available, including bias-reduced procedures [5], [20], [21], robust techniques [8], and algorithms for small samples [6]. In this study, because of its optimal asymptotic properties [22], the maximum likelihood (ML) estimation is employed.

Let $z_1, z_2, \ldots, z_n$ be a random sample of size $n$ from $G^0(\alpha, \gamma, L)$. The maximum likelihood estimators for the parameters $\alpha$ and $\gamma$, namely $\hat{\alpha}$ and $\hat{\gamma}$, are the solution of the following system of non-linear equations [10], [11]:

$$\psi^0(L - \hat{\alpha}) - \psi^0(-\hat{\alpha}) - \log(\hat{\gamma}) + \frac{1}{\hat{\gamma}} \sum_{i=1}^{n} \log(\hat{\gamma} + L z_i) = 0,$$

$$-\hat{\alpha} + \frac{\hat{\alpha} - L}{\hat{\gamma}} \sum_{i=1}^{n} (\hat{\gamma} + L z_i)^{-1} = 0,$$

where $\psi^0(\cdot)$ is the digamma function. In general, the above system of equations does not possess a closed form solution and numerical optimization methods are required. We use the BFGS method, which is regarded as an accurate technique [23].

Next section presents several contrast measures between $G^0$ distributions.

III. CONTRAST MEASURES FOR THE $G^0$ LAW

Contrast analysis often addresses the problem of quantifying how distinguishable two image regions are from each other. An image can be interpreted as a set of regions described by possibly different probability laws. Many statistical approaches have been developed to reach this goal, leading to the difficult problem of specifying an accurate, expressive and tractable metrics. In general terms, these methods can be either parametric or nonparametric.

In a previous work, we proposed a class of parametric measures based on stochastic distances. Advancing that study, among the existing nonparametric methods, we now include an approach based on the Kolmogorov-Smirnov test. This method is reported to possess good discriminatory properties [24] and has found applications in several instances [9], [12].

In the following, we briefly describe the referred parametric methods. Moreover, we adapt the Kolmogorov-Smirnov test for $G^0$ distributed data. We assume that $Z'$ and $Z''$ are random variables with cumulative distribution functions given by $F_{\theta_1}$ and $F_{\theta_2}$, respectively. These distributions are defined over the same probability space, equipped with densities $f_{Z'}(z; \theta_1)$ and $f_{Z''}(z; \theta_2)$, respectively, where $\theta_1$ and $\theta_2$ are parameter vectors; and they share a common support $I \subset \mathbb{R}$.

Section III-A presents those contrast measures derived from stochastic distances and an example of their application to real SAR data. Section III-B discusses the use of the Kolmogorov-Smirnov distance for the $G^0$ model.

A. Contrast Measures Based on Stochastic Distances

Information theoretical tools collectively known as divergence measures offer entropy based methods for contrasting stochastic distributions [2]. Often not rigorously metrics, since the triangle inequality does not necessarily holds, divergence measures are mathematically suitable tools for comparing the distribution of random variables [25]. Divergence measures were submitted to a systematic and comprehensive treatment and, as a result, the class of $(h, \phi)$-divergences was proposed [22].

The $(h, \phi)$-divergence between $f_{Z'}$ and $f_{Z''}$ is defined by

$$D^h_{\phi}(Z', Z'') = h \left( \int \phi \left( \frac{f_{Z'}(z; \theta_1)}{f_{Z''}(z; \theta_2)} \right) f_{Z''}(z; \theta_2) dz \right),$$

where $\phi: (0, \infty) \to [0, \infty)$ is a convex function, $h: (0, \infty) \to [0, \infty)$ is a strictly increasing function with $h(0) = 0$, and indeterminate forms are assigned value zero.

By a judicious choice of $h$ and $\phi$, some well-known divergence measures arise. Table I shows a selection of functions $h$ and $\phi$ that lead to notable distance measures. This set of distance measures was analyzed and applied to speckled data modeled by $G^0$ law in [10], [11]. Since the Kullback-Leibler divergence is not a symmetric measure, it was symmetrized [26] and the resulting measure is termed “distance”.

In the following we provide integral expressions for the discussed $(h, \phi)$-distances. For simplicity, we suppress the explicit dependence on $x$ and on the support $I$.

(i) The Kullback-Leibler distance:

$$d_{KL}(Z', Z'') = \frac{1}{2} \int (f_{Z'} - f_{Z''}) \log \left( \frac{f_{Z'}}{f_{Z''}} \right).$$

(ii) The triangular distance:

$$d_T(Z', Z'') = \int \frac{(f_{Z'} - f_{Z''})^2}{f_{Z'} + f_{Z''}}.$$

(iii) The Bhattacharyya distance:

$$d_B(Z', Z'') = -\log \left( \int \sqrt{f_{Z'} f_{Z''}} \right).$$

(iv) The arithmetic-geometric distance:

$$d_{AG}(Z', Z'') = \frac{1}{2} \int (f_{Z'} + f_{Z''}) \log \left( \frac{f_{Z'} + f_{Z''}}{2\sqrt{f_{Z'} f_{Z''}}} \right).$$

(v) The Rényi distance of order $\beta \in (0, 1)$:

$$d^\beta_R(Z', Z'') = \frac{1}{\beta - 1} \log \left( \frac{\int f_{Z'}^\beta f_{Z''}^{1-\beta} + \int f_{Z'}^{1-\beta} f_{Z''}^\beta}{2} \right).$$

(vi) The Jensen-Shannon distance:

$$d_{JS}(Z', Z'') = \frac{1}{2} \int f_{Z'} \log \left( \frac{2f_{Z'}}{f_{Z'} + f_{Z''}} \right) + f_{Z''} \log \left( \frac{2f_{Z''}}{f_{Z'} + f_{Z''}} \right).$$

(vii) The Hellinger distance:

$$d_H(Z', Z'') = 1 - \sqrt{f_{Z'} f_{Z''}} = 1 - \exp \left( -\frac{1}{2} d^1_R(Z', Z'') \right).$$

(viii) The harmonic-mean distance:

$$d_{HM}(Z', Z'') = -\log \left( \frac{2f_{Z'} f_{Z''}}{f_{Z'} + f_{Z''}} \right) = -\log \left( 1 - \frac{d_T(Z', Z'')}{2} \right).$$
Several convergence properties of the \((h, \phi)\)-divergences were established in [22]. Under the regularity conditions discussed in [22, p. 380], if \(\hat{\theta}_1 = \theta_2\) then, as \(m, n \to \infty\),

\[
\frac{2mn}{m + n} \frac{D^h_\phi(\hat{\theta}_1, \hat{\theta}_2)}{h^{(0)}(0)\phi''(1)} \xrightarrow{\mathcal{D}} \chi^2_M,
\]

is asymptotically chi-square distributed with \(M\) degrees of freedom, where \(\hat{\theta}_1 = (\hat{\theta}_{11}, \ldots, \hat{\theta}_{1M})\) and \(\hat{\theta}_2 = (\hat{\theta}_{21}, \ldots, \hat{\theta}_{2M})\) are the ML estimators of \(\theta_1\) and \(\theta_2\) based on independent samples of sizes \(m\) and \(n\), respectively [22].

Thus, when (i) considering the definition of the distances in terms of the \(h\) and \(\phi\) functions, and (ii) applying the results on the convergence in distribution of the \((h, \phi)\)-measures to \(\chi^2_M\) [22], the lemma below is proved.

**Lemma 1.** Let the regularity conditions proposed in [22, p. 380] hold. If \(\frac{m}{m+n} \xrightarrow{\mathcal{D}} \lambda \in (0, 1)\) and \(\theta_1 = \theta_2\), then

\[
\frac{2mn}{m + n} \frac{d^h_\phi(\hat{\theta}_1, \hat{\theta}_2)}{h^{(0)}(0)\phi''(1)} \xrightarrow{\mathcal{D}} \chi^2_M,
\]

where \(\xrightarrow{\mathcal{D}}\) denotes convergence in distribution.

Based on Lemma 1, statistical hypothesis tests for the null hypothesis \(\theta_1 = \theta_2\) can be derived. In particular, the following statistic is considered:

\[
S^h_\phi(\hat{\theta}_1, \hat{\theta}_2) = \frac{2mnv}{m + n} \frac{d^h_\phi(\hat{\theta}_1, \hat{\theta}_2)}{h^{(0)}(0)\phi''(1)},
\]

where \(v = 1/ (h^{(0)}(0)\phi''(1))\) is a constant that depends on the chosen distance. Table II lists the values of \(v\) for each distance. We are now in position to state the following result.

**Proposition 1:** Let \(m\) and \(n\) assume large values and \(S^h_\phi(\hat{\theta}_1, \hat{\theta}_2) = s\), then the null hypothesis \(\theta_1 = \theta_2\) can be rejected at a level \(\eta\) if \(\Pr(\chi^2_M > s) \leq \eta\).

### TABLE I

<table>
<thead>
<tr>
<th>((h, \phi))-distance</th>
<th>(h(y))</th>
<th>(\phi(x))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kullback-Leibler</td>
<td>(\frac{y}{2})</td>
<td>((x - 1)\log(x) - \frac{(x - 1)^2}{2})</td>
</tr>
<tr>
<td>Triangular</td>
<td>(-\log(-y + 1), 0 \leq y &lt; 1)</td>
<td>(-\frac{2}{x + 2} + \frac{1}{x^2})</td>
</tr>
<tr>
<td>Bhattacharyya</td>
<td>(-\log(y), 0 \leq y &lt; 2)</td>
<td>(-\frac{4}{x + 4} + \frac{1}{x^2})</td>
</tr>
<tr>
<td>Arithmetic-geometric</td>
<td>(-\log((\beta - 1)y + 1), 0 \leq y &lt; \frac{1}{\beta - 1})</td>
<td>(-\frac{2}{x + 2\beta} + \frac{1}{x^2})</td>
</tr>
<tr>
<td>Rényi (order (\beta \in (0, 1)))</td>
<td>(-\log\left(-\frac{x}{\beta} + 1\right), 0 \leq y &lt; 2)</td>
<td>(-\frac{2}{x + 2\beta} + \frac{1}{x^2})</td>
</tr>
<tr>
<td>Jensen-Shannon</td>
<td>(-\log\left(-\frac{x}{\beta} + 1\right), 0 \leq y &lt; 2)</td>
<td>(-\frac{2}{x + 2\beta} + \frac{1}{x^2})</td>
</tr>
<tr>
<td>Hellinger</td>
<td>(-\log\left(-\frac{x}{\beta} + 1\right), 0 \leq y &lt; 2)</td>
<td>(-\frac{2}{x + 2\beta} + \frac{1}{x^2})</td>
</tr>
<tr>
<td>Harmonic-Mean</td>
<td>(-\log\left(-\frac{x}{\beta} + 1\right), 0 \leq y &lt; 2)</td>
<td>(-\frac{2}{x + 2\beta} + \frac{1}{x^2})</td>
</tr>
</tbody>
</table>

### TABLE II

<table>
<thead>
<tr>
<th>Distance</th>
<th>(v)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kullback-Leibler</td>
<td>1</td>
</tr>
<tr>
<td>Triangular</td>
<td>1</td>
</tr>
<tr>
<td>Bhattacharyya</td>
<td>1</td>
</tr>
<tr>
<td>Arithmetic-geometric</td>
<td>4</td>
</tr>
<tr>
<td>Rényi (order (\beta))</td>
<td>(1/\beta)</td>
</tr>
<tr>
<td>Jensen-Shannon</td>
<td>4</td>
</tr>
<tr>
<td>Hellinger</td>
<td>4</td>
</tr>
<tr>
<td>Harmonic-Mean</td>
<td>2</td>
</tr>
</tbody>
</table>

In terms of image analysis, this proposition offers a method to statistically refute the hypothesis that two samples obtained in different regions can be described by the same distribution. Such assessment is relevant, for instance, when choosing input training samples in supervised learning procedures or when proposing similarity-based techniques such as segmentation.

Figure 3 presents a SAR image for which the estimated number of looks is 3.2. It was obtained by the E-SAR sensor over surroundings of Munich, Germany [27]. The area exhibits three distinct types of target roughness: (i) homogeneous (corresponding to pasture), (ii) heterogeneous (forest), and (iii) extremely heterogeneous (urban areas). Samples were selected and submitted to statistical analysis, after parameters estimation by maximum likelihood methods.

Maximum likelihood estimates for \(G^0\) distributed data are often difficult to obtain due to numerical instability issues [5]. This problem was previously reported in [6], and estimate censoring was proposed as a procedure to circumvent this situation.

Table III shows the estimates in each of these samples, as well as their size (# pixels) and the number of disjoint \(7 \times 7\) pixels blocks samples (# parts). Those blocks were used to estimate the test sizes and powers.

### TABLE III

<table>
<thead>
<tr>
<th>Regions</th>
<th>(\alpha)</th>
<th>(\gamma)</th>
<th># pixels</th>
<th># parts</th>
</tr>
</thead>
<tbody>
<tr>
<td>pasture-1</td>
<td>13.702</td>
<td>392.39</td>
<td>1235</td>
<td>25</td>
</tr>
<tr>
<td>pasture-2</td>
<td>-12.698</td>
<td>803.20</td>
<td>1216</td>
<td>24</td>
</tr>
<tr>
<td>pasture-3</td>
<td>-11.304</td>
<td>1622.92</td>
<td>1602</td>
<td>40</td>
</tr>
<tr>
<td>forest</td>
<td>-9.339</td>
<td>661.183</td>
<td>1606</td>
<td>32</td>
</tr>
<tr>
<td>urban-1</td>
<td>-0.759</td>
<td>1484.13</td>
<td>2005</td>
<td>40</td>
</tr>
<tr>
<td>urban-2</td>
<td>-0.388</td>
<td>1101.83</td>
<td>3481</td>
<td>71</td>
</tr>
<tr>
<td>urban-3</td>
<td>-1.079</td>
<td>555.83</td>
<td>4657</td>
<td>95</td>
</tr>
</tbody>
</table>

Table IV presents the observed rejection rates of samples from the discussed region at 1% significance level, i.e., the test size or Type I error. The results show that \(S_{KL}\), \(S_T\), and \(S_{AG}\) tests have excellent performance with respect to this criterion, and that the hardest situation is related to urban areas. Moreover, the \(S_H\), \(S_B\), \(S_{HM}\) and \(S_R\) test present instability with the worst cases in urban regions.

Table V shows the test powers at 1% nominal levels, i.e., the inability to reject samples from different types, or Type II error. The best performances are observed when confronting...
area whose distributions differ more significantly. In general, the measure linked to the $S_{AG}$ statistic presented the best performance among those selected. Although the other proposed tests were surpassed, their results were comparable to the best ones.

In the remainder of this text only the classical distances, namely the ones derived from the Kullback-Leibler ($S_{KL}$) and Bhattacharyya ($S_{B}$) tests, and the best ones derived from $(h\phi)$-divergence tests, i.e., triangular ($S_{T}$) and arithmetic-geometric ($S_{AG}$) will be compared with the one derived from the Kolmogorov-Smirnov test.

### Table IV

<table>
<thead>
<tr>
<th>Regions</th>
<th>$S_{KL}$</th>
<th>$S_{B}$</th>
<th>$S_{T}$</th>
<th>$S_{AG}$</th>
<th>$S_{KL}$</th>
<th>$S_{B}$</th>
<th>$S_{T}$</th>
<th>$S_{AG}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>pasture-1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>pasture-2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>pasture-3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.36</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.36</td>
<td>0.00</td>
</tr>
<tr>
<td>forest</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>urban-1</td>
<td>0.00</td>
<td>0.13</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.13</td>
<td>0.00</td>
</tr>
<tr>
<td>urban-2</td>
<td>0.00</td>
<td>99.96</td>
<td>0.00</td>
<td>99.96</td>
<td>0.00</td>
<td>99.20</td>
<td>0.00</td>
<td>100.00</td>
</tr>
<tr>
<td>urban-3</td>
<td>4.23</td>
<td>3.23</td>
<td>1.97</td>
<td>3.56</td>
<td>2.73</td>
<td>2.15</td>
<td>5.64</td>
<td>15.03</td>
</tr>
</tbody>
</table>

### Table V

Rejection rates of $(h, \phi)$-divergence tests under $H_1: (\alpha_1, \gamma_1) \neq (\alpha_2, \gamma_2)$ with $\mu_1 = \mu_2$ at 1% nominal level.

<table>
<thead>
<tr>
<th>Regions</th>
<th>$S_{KL}$</th>
<th>$S_{B}$</th>
<th>$S_{T}$</th>
<th>$S_{AG}$</th>
<th>$S_{KL}$</th>
<th>$S_{B}$</th>
<th>$S_{T}$</th>
<th>$S_{AG}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>pasture-1</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>pasture-2</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>pasture-3</td>
<td>94.84</td>
<td>90.48</td>
<td>87.50</td>
<td>90.48</td>
<td>92.26</td>
<td>88.69</td>
<td>95.24</td>
<td>92.26</td>
</tr>
<tr>
<td>forest</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>urban-1</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>urban-2</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>urban-3</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>urban-1</td>
<td>98.70</td>
<td>99.70</td>
<td>98.10</td>
<td>99.70</td>
<td>98.50</td>
<td>99.20</td>
<td>98.40</td>
<td>99.80</td>
</tr>
<tr>
<td>urban-2</td>
<td>98.03</td>
<td>100</td>
<td>99.49</td>
<td>100</td>
<td>99.27</td>
<td>100</td>
<td>97.30</td>
<td>100</td>
</tr>
</tbody>
</table>

**B. Contrast measure based on the Kolmogorov-Smirnov test**

Let $\{z'_1, z'_2, \ldots, z'_m\}$ and $\{z''_1, z''_2, \ldots, z''_n\}$ be two independent random samples of sizes $m$ and $n$ drawn from population variables $Z'$ and $Z''$, respectively. We aim to test the null hypothesis $H_0: F_{a_1} = F_{a_2}$ against the alternative hypothesis $H_1: F_{a_1} \neq F_{a_2}$. Additionally, admit that $\{z'_i: i = 1, \ldots, m\}$ and $\{z''_i: i = 1, \ldots, n\}$ are the order statistics associated to the original samples, respectively. The empirical distributions of these order statistics are

$$S_m(z) = \begin{cases} 0, & \text{if } z < z'_1, \\ k/m, & \text{if } z'_k \leq z \leq z'_{(k+1)}, \\ 1, & \text{if } z > z'_m, \end{cases} \quad (4)$$

and

$$T_n(z) = \begin{cases} 0, & \text{if } z < z''_1, \\ k/n, & \text{if } z''_k \leq z \leq z''_{(k+1)}, \\ 1, & \text{if } z > z''_n, \end{cases} \quad (5)$$

for $Z'$ and $Z''$, respectively.

The statistics in Equations (4) and (5) can be used to form the Kolmogorov-Smirnov empirical distance defined as [28]:

$$d_{m,n} = \max_{-\infty < z < \infty} |S_m(z) - T_n(z)|.$$
Smirnov [28] provided a useful result on the asymptotics of the Kolmogorov-Smirnov distance which is stated as follows. If \( m, n \to \infty \) in such a way that \( mn/n \) remains constant, then
\[
\lim_{m,n \to \infty} \Pr \left( \frac{mn}{m+n} d_{m,n} \leq \delta \right) = L(\delta),
\]
where \( L(\delta) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} \exp \left( -2i^2\delta^2 \right) \) and \( \delta \) is a critical value.

Based on this limiting behavior under \( H_0 \), the Kolmogorov-Smirnov test statistic for \( G^0 \)-distributed data is given by:
\[
S_{KS}(Z', Z'') = \sqrt{\frac{mn}{m+n}} \hat{d}_{KS}(Z', Z''),
\]
where \( \hat{d}_{KS}(Z', Z'') = d_{m,n} \). Gilvenko-Cantelli theorem ensures that this statistic test is consistent for the alternative hypothesis [29]. Therefore, in view of the above, the following result holds.

**Proposition 2:** Let \( m \) and \( n \) assume large values, then the null hypothesis \( \tilde{H}_0 \) of \( \hat{F}_n \) can be rejected at a level of \( \eta \) if
\[
\Pr(S_{KS}(Z', Z'') > \delta) = 1 - L(\delta) \leq \eta.
\]

Next section presents a comprehensive comparison of the distances derived in section III-A and the distances computed in section III-B using both the pure \( G^0 \) and a contaminated model.

**IV. ANALYSIS WITH SIMULATED DATA**

Synthetic SAR images modelled with \( G^0(\alpha, \gamma, L) \) laws were employed to assess the discussed parametric and nonparametric test statistics. The parameters employed in the simulation were \( \alpha \in \{-1.5, -3, -5\}, \mu \in \{1, 2, 3\} \) and \( L \in \{1, 8\} \), where \( \mu = -\gamma/(1 + \alpha) \) is the mean value of the \( G^0 \) distribution; cf. Equation (3). The quantity \( \mu \) is directly related to the image brightness. We checked the tests at the 1% and 10% significance levels using square windows of sizes 7 × 7, 9 × 9, and 11 × 11 pixels (\( N \in \{49, 81, 121\} \)).

The empirical size and power of the proposed tests were sought as a means to measure and compare their performances. Monte Carlo experiments under three scenarios were designed to obtain empirical data. Considering two image regions specified by the parameter vectors \((\alpha_1, \mu_1, L)\) and \((\alpha_2, \mu_2, L)\), the following situations were set: (i) \( \alpha_1 = \alpha_2, \mu_1 = \mu_2, \) (ii) \( \alpha_1 = \alpha_2, \mu_1 \neq \mu_2, \) and (iii) \( \alpha_1 > \alpha_2, \mu_1 > \mu_2. \) Whereas the first situation corresponds to image regions of same roughness and brightness, scenario (i) accounts for images of same roughness but with different brightness. Images of both distinct roughness and brightness are captured by scenario (iii).

In practice, it is quite hard to guarantee that all observations in a sample come from random variables with the same distribution. When collecting samples by visual inspection, frequently the onset of some type of contamination is experienced. For example, a few values could come from a different distribution other than the one assumed. One of the most common sources of contamination in SAR imagery are connected to the double bounce phenomenon: a few pixels exhibit a very high return caused by either natural or man-made targets that send back most of the energy they receive. This is the case of flooded forests and urban areas. The presence of a single pixel suffering from double bounce may produce meaningless estimates and test statistics, unless they exhibit some robustness with respect to this kind of contamination. Figure 4 shows an example of the double bounce effect on an homogeneous field.

![Fig. 4. A corner reflector producing double bounce on a homogeneous background](image)

In order to assess the robustness of the considered tests, we assume contamination in a parametric way. In other words, it is admitted that the data in a sample come either from the hypothesized distribution \( G^0 \) with probability \( 1 - \epsilon \), or from a scaled version of the same law with probability \( \epsilon \). The considered scaled distribution is such that its mean value is one hundred times larger than the mean of the original distribution. Let \( A \) be the event "presence of outlier", which has \( \Pr(A) = \epsilon \). By the total probability rule, the cumulative distribution function of the return is
\[
F_Z(t) = \Pr \left( \{ Z \leq t \} \right) = \Pr \left( \{ Z \leq t \} \cap A \right) + \Pr \left( \{ Z \leq t \} \cap A^c \right)
\]
\[
= \Pr(A^c) \Pr \left( Z \leq t \mid A^c \right) + \Pr(A) \Pr \left( Z \leq t \mid A \right)
\]
\[
= (1 - \epsilon) F_{G^0(\alpha, \gamma, L)}(t) + \epsilon F_{\hat{G}(\alpha, 10\gamma, L)}(t).
\]

Following [14], [15], we chose to assess the situations without contamination \( \epsilon = 0 \) and with mild levels of contamination \( \epsilon \in \{10^{-4}, 10^{-3}, 5 \times 10^{-3}\} \).

The simulations were based on 5500 replications, and the only valid instances were those where the estimated parameter \( \alpha \) was in the range \([10\alpha, \alpha/20]\), due to censoring.

Table VI presents the simulation results under scenario (i). With the exception of \( S_T \), the estimated significance level is above the nominal value. Overall, the \( S_T \) measure is the most efficient regarding the test size, since its Type I error is the closest to the nominal values. The measure based on the Kolmogorov-Smirnov test \( S_{KS} \) furnished one of the lowest estimated values. This can be explained by the fact that it aims at testing equality between distributions.

In the 54 examined situations, the tests had empirical sizes not greater than the nominal level as follows: triangular in 100% of the cases, Kolmogorov-Smirnov in 99.07%, Bhatacharyya in 67.60%, Kullback-Leibler in 49.07%, and, finally,
arithmetico-geometric in 37.03%. It is noteworthy that the two most commonly employed distances, namely the Kullback-Leibler and the Bhattacharyya distances, presented poor performance when used as test statistics. These values were obtained with 2118 minimum, 4692 mean, and 5500 maximum valid replications after censoring.

Tables VII and VIII present the results regarding scenarios (ii) and (iii), respectively. Table VII was obtained with 3449 minimum, 5037 mean, and 5500 maximum valid replications after censoring, while those values are 2572, 4630, and 5500 for Table VIII.

The simulation results show that the $S_{AG}$ test outperforms other techniques. Both tables show that the power increases both with the sample size $N$ and with the number of looks $L$. Moreover, the observed powers appear in all but two situations sorted as $\text{Power}(S_{AG}) \geq \text{Power}(S_{KL}) \geq \text{Power}(S_{B}) \geq \text{Power}(S_{S})$.

From Table VII we conclude that under the alternative hypothesis of only different mean brightness, the parametric measures behave substantially better than $S_{KL}$. Table VIII shows that there is a substantial loss of power when dealing with the alternative hypothesis $(\alpha_1, \mu_1) \neq (\alpha_2, \mu_2)$ such that $\mu_1 > \mu_2$ and $\alpha_1 > \alpha_2$, since it is a less tractable situation as shown in [10], [11]. However, unlike the nonparametric measure $S_{KS}$, parametric tests based on distances have proven to be increasingly more accurate with (i) the number of looks, (ii) the sample size, and (iii) the nominal level.

Based on the results presented in Tables VI and VIII, we consider reasonable to restrict the forthcoming detailed robustness study to the best statistics, namely, $S_{T}$ and $S_{AG}$, and to the one that is expected to exhibit robustness, i.e., to the one based on the Kolmogorov-Smirnov test $S_{KS}$. For these measures, Figures 5 and 6 show the behavior of the test size in the presence of a fixed level $\varepsilon$ of contamination.

The graphs suggest that the empirical test size grows with the window size when there is contamination, i.e., for $\varepsilon \in \{-4 \times 10^{-4}, -5 \times 10^{-4} \}$. In most cases, the test size based in $S_{T}$ is the most robust for homogeneous regions. We also note that the empirical test sizes of $(h, \delta)$-distance based tests for $\alpha = -1.5$ is greater than one observed when $\alpha \in \{-3, -5\}$; and this behavior is more pronounced as $\varepsilon$ increases.

Surprisingly, the Kolmogorov-Smirnov test tends to be less robust than the parametric tests when the contamination increases. Moreover, the similarity of the test size curves indicates that the image roughness and brightness do not significantly affect the estimated test sizes.
This paper presented a comparison among parametric tests based on stochastic distances and the Kolmogorov-Smirnov test. The assessment was made with Monte Carlo experiments varying the parameters of the $G^3$ distribution, regarded to an universal model for speckled data, and the level of a plausible contamination. We also presented compact formulas for the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.

We show numerical evidence that the test based on the triangular distance $S_T$ has, in general, smaller empirical Type I errors than the test based on the Kolmogorov-Smirnov distance $S_{KS}$. Using a parametric and plausible contamination model, we illustrate that when the chances of observing aberrant data are higher, the test based on $S_{KS}$ has its performance diminished compared to the Kolmogorov-Smirnov test.

For the Kolmogorov-Smirnov contrast measure. In addition, we consider a robustness study to assess the accuracy of the measures under examination.
For each sample size, all considered scenario situations were ordered and indexed according to $i$. The parameter values: (a) $0$, (b) $10^{-4}$, (c) $10^{-3}$ and (d) $5 \times 10^{-3}$. For each sample size, all considered scenario situations were ordered and indexed according to $i$. The index $i$ was employed as the abscissa values.

**Fig. 6.** Test size under the model in Equation 6 for nominal level 10% and $\varepsilon$ parameter values: (a) 0, (b) $10^{-4}$, (c) $10^{-3}$ and (d) $5 \times 10^{-3}$. (a), (b), (c), (d)