

CHILEAN JOURNAL OF STATISTICS

Edited by Víctor Leiva and Carolina Marchant

A free open access journal indexed by



Web of
Science
Group

Scopus®



ELSEVIER

Volume 12 Number 1
April 2021

ISSN: 0718-7912 (print)

ISSN: 0718-7920 (online)

Published by the
Chilean Statistical Society

SOCH E 
SOCIEDAD CHILENA DE ESTADÍSTICA

AIMS

The Chilean Journal of Statistics (ChJS) is an official publication of the Chilean Statistical Society (www.soche.cl). The ChJS takes the place of *Revista de la Sociedad Chilena de Estadística*, which was published from 1984 to 2000.

The ChJS covers a broad range of topics in statistics, as well as in artificial intelligence, big data, data science, and machine learning, focused mainly on research articles. However, review, survey, and teaching papers, as well as material for statistical discussion, could be also published exceptionally. Each paper published in the ChJS must consider, in addition to its theoretical and/or methodological novelty, simulations for validating its novel theoretical and/or methodological proposal, as well as an illustration/application with real data.

The ChJS editorial board plans to publish one volume per year, with two issues in each volume. On some occasions, certain events or topics may be published in one or more special issues prepared by a guest editor.

EDITORS-IN-CHIEF

Víctor Leiva
Carolina Marchant

Pontificia Universidad Católica de Valparaíso, Chile
Universidad Católica del Maule, Chile

EDITORS

Héctor Allende Cid	<i>Pontificia Universidad Católica de Valparaíso, Chile</i>
Danilo Alvares	<i>Pontificia Universidad Católica de Chile</i>
José M. Angulo	<i>Universidad de Granada, Spain</i>
Robert G. Aykroyd	<i>University of Leeds, UK</i>
Narayanaswamy Balakrishnan	<i>McMaster University, Canada</i>
Michelli Barros	<i>Universidade Federal de Campina Grande, Brazil</i>
Carmen Batanero	<i>Universidad de Granada, Spain</i>
Ionut Bebu	<i>The George Washington University, US</i>
Marcelo Bourguignon	<i>Universidade Federal do Rio Grande do Norte, Brazil</i>
Márcia Branco	<i>Universidade de São Paulo, Brazil</i>
Oscar Bustos	<i>Universidad Nacional de Córdoba, Argentina</i>
Luis M. Castro	<i>Pontificia Universidad Católica de Chile</i>
George Christakos	<i>San Diego State University, US</i>
Enrico Colosimo	<i>Universidade Federal de Minas Gerais, Brazil</i>
Gauss Cordeiro	<i>Universidade Federal de Pernambuco, Brazil</i>
Francisco Cribari-Neto	<i>Universidade Federal de Pernambuco, Brazil</i>
Francisco Cysneiros	<i>Universidade Federal de Pernambuco, Brazil</i>
Mário de Castro	<i>Universidade de São Paulo, São Carlos, Brazil</i>
José A. Díaz-García	<i>Universidad Autónoma Agraria Antonio Narro, Mexico</i>
Raul Fierro	<i>Universidad de Valparaíso, Chile</i>
Jorge Figueroa-Zúñiga	<i>Universidad de Concepción, Chile</i>
Isabel Fraga	<i>Universidade de Lisboa, Portugal</i>
Manuel Galea	<i>Pontificia Universidad Católica de Chile</i>
Diego Gallardo	<i>Universidad de Atacama, Chile</i>
Christian Genest	<i>McGill University, Canada</i>
Viviana Giampaoli	<i>Universidade de São Paulo, Brazil</i>
Marc G. Genton	<i>King Abdullah University of Science and Technology, Saudi Arabia</i>
Patricia Giménez	<i>Universidad Nacional de Mar del Plata, Argentina</i>
Hector Gómez	<i>Universidad de Antofagasta, Chile</i>
Yolanda Gómez	<i>Universidad de Atacama, Chile</i>
Emilio Gómez-Déniz	<i>Universidad de Las Palmas de Gran Canaria, Spain</i>
Daniel Griffith	<i>University of Texas at Dallas, US</i>
Eduardo Gutiérrez-Peña	<i>Universidad Nacional Autónoma de México</i>
Nikolai Kolev	<i>Universidade de São Paulo, Brazil</i>
Eduardo Lalla	<i>University of Twente, Netherlands</i>
Shuangzhe Liu	<i>University of Canberra, Australia</i>
Jesús López-Fidalgo	<i>Universidad de Navarra, Spain</i>
Liliana López-Kleine	<i>Universidad Nacional de Colombia</i>
Rosângela H. Loschi	<i>Universidade Federal de Minas Gerais, Brazil</i>
Manuel Mendoza	<i>Instituto Tecnológico Autónomo de México</i>
Orietta Nicolis	<i>Universidad Andrés Bello, Chile</i>
Ana B. Nieto	<i>Universidad de Salamanca, Spain</i>
Teresa Oliveira	<i>Universidade Aberta, Portugal</i>
Felipe Osorio	<i>Universidad Técnica Federico Santa María, Chile</i>
Carlos D. Paulino	<i>Instituto Superior Técnico, Portugal</i>
Fernando Quintana	<i>Pontificia Universidad Católica de Chile</i>
Nalini Ravishanker	<i>University of Connecticut, US</i>
Fabrizio Ruggeri	<i>Consiglio Nazionale delle Ricerche, Italy</i>
José M. Sarabia	<i>Universidad de Cantabria, Spain</i>
Helton Saulo	<i>Universidade de Brasília, Brazil</i>
Pranab K. Sen	<i>University of North Carolina at Chapel Hill, US</i>
Giovani Silva	<i>Universidade de Lisboa, Portugal</i>
Julio Singer	<i>Universidade de São Paulo, Brazil</i>
Milan Stehlik	<i>Johannes Kepler University, Austria</i>
Alejandra Tapia	<i>Universidad Católica del Maule, Chile</i>
M. Dolores Ugarte	<i>Universidad Pública de Navarra, Spain</i>

Chilean Journal of Statistics

VOLUME 12, NUMBER 1

APRIL 2021

CONTENTS

Carolina Marchant and Víctor Leiva <i>Chilean Journal of Statistics: A forum for the Americas and the World in COVID-19 pandemic</i>	1
Ruth Burkhalter and Yuhlong Lio <i>Bootstrap control charts for the generalized Pareto distribution percentiles</i>	3
Carlos López-Vázquez, Andrómaca Tasistro, and Esther Hochsztain <i>Exact tables for the Friedman rank test: Case with ties</i>	23
Lucas de Oliveira Ferreira de Sales, André Luís Santos de Pinho, Francisco Moisés Cândido de Medeiros, and Marcelo Bourguignon <i>Control chart for monitoring the mean in symmetric data</i>	37
Agatha S. Rodrigues, Vinicius F. Calsavara, Eduardo Bertolli, Stela V. Peres, and Vera L. D. Tomazella <i>Bayesian long-term survival model including a frailty term: Application to melanoma datas</i>	53
Abraão D.C. Nascimento, Kássio F. Silva, and Alejandro C. Frery <i>Distance-based edge detection on synthetic aperture radar imagery</i>	71
Malinda Coa and Ernesto Ponsot <i>Alternatives to the logit model in the situation of factor levels aggregation in binomial responses</i>	83
Francisco J. Ariza-Hernandez and Eduardo Gutiérrez-Peña <i>Bayesian analysis of an item response model with an AEP-based link function</i>	103

SPATIAL STATISTICS AND IMAGE MODELING
RESEARCH PAPER

Distance-based edge detection on synthetic aperture radar imagery

ABRAÃO D.C. NASCIMENTO^{1,*}, KÁSSIO F. SILVA¹, AND ALEJANDRO C. FRERY²

¹DEPARTAMENTO DE ESTATÍSTICA, UNIVERSIDADE FEDERAL DE PERNAMBUCO, AV.
PROFESSOR MORAES REGO, 1235, RECIFE 50670-901, BRAZIL

²SCHOOL OF MATHEMATICS AND STATISTICS, VICTORIA UNIVERSITY OF WELLINGTON,
NEW ZEALAND

(Received: 01 March 2021 · Accepted in final form: 22 April 2021)

Abstract

Synthetic aperture radar is an efficient remote sensing tool by producing high spatial resolution images. But, synthetic aperture radar data suffer speckle noise effect that difficult their processing (for example, making boundary detection). We propose and assess edge detectors for synthetic aperture radar imagery based on stochastic distances between models. These edge detectors stem from generalized divergences with good asymptotic properties. Results reveal that divergence-based detectors can outperform the likelihood-based counterpart.

Keywords: Edge detection · \mathcal{G}_I^0 model · Synthetic aperture radar systems · Speckled data · Stochastic distance.

Mathematics Subject Classification: Primary 62B10 · Secondary 68U10.

1. INTRODUCTION

Because of its all-time, all-weather, high-penetration, and high-resolution imaging capability on a global scale, synthetic aperture radar (SAR) has become an essential tool for land survey, resource mapping, environmental monitoring, disaster rescue, and national security. SAR systems have progressed from low to high resolution, single polarization to full polarization, and single frequency to multifrequency. SAR images can be analyzed using a variety of techniques. There are primarily two types of methods depending on their theoretical foundations: electromagnetic (EM) physics methods based on Maxwell's equations (Kong, 1990), and statistical methods which focus on the image data.

Due to the high complexity of the EM approach, both theoretically and computationally, only simplified or empirical models for specific scenarios can be created. The statistical approach is focused on the relationships between pixel values and their distributional characteristics.

*Corresponding author. Email: abraao@de.ufpe.br

SAR image statistical analysis can be traced back to the 1950s. The first statistical study was focused on SAR clutter in the ocean. Since early radar images had poor resolution, the Rayleigh speckle model was developed under the assumption of the Central Limit Theorem, resulting in the Rayleigh distribution for the amplitude of radar echoes (Ward et al, 2006).

The Rayleigh model, however, became less reliable as the spatial resolution of SAR images improved: smaller areas comprise less elements in the summation, making the large sample assumption questionable in many cases.

Ward (1981) proposed the multiplicative model in 1981, a turning point in statistical SAR data description. Such an approach generalizes the Rayleigh model, and bridges the EM and statistical approaches (Yue et al, 2020, 2021)

The central problem to be described is the presence of an interference pattern, common to all images obtained with coherent illumination, called *speckle*. Although deterministic, the precise knowledge of speckle amounts to specifying the EM characteristics of each scattering element within a resolution cell. This is possible when there are a few well-known and simple backscatterers as, for instance, a single small sphere or a few dipoles (Sant'Anna et al, 2008). The multiplicative model is an adequate approach when the number of backscatterers or their properties are unknown.

The models that arise from the multiplicative description are neither Gaussian nor additive. Classical image processing techniques are, at best, sub-optimal in such scenario. Therefore, SAR imagery processing requires specialized models and methodologies. To that end, the use of Information Statistical Theory measures combined with the multiplicative modeling approach has been successfully adopted for treating SAR images.

Edge detection is one of the fundamental image processing techniques. Gambini et al (2006) proposed a method that relies on comparing two samples for estimating the position of the edge along a thin strip of data. Gambini et al (2008) compared five strategies based on SAR data or on estimates of the target roughness. Wei and Feng (2015), assuming a gamma model, derived a detector with low false alarm rate, but its performance strongly depends on the settings. Giron et al (2012) used a nonparametric approach with good results but, again, the performance is affected by the underlying distribution of the data.

In this paper, we assume that intensity SAR data follow the \mathcal{G}_I^0 model. This distribution is recognized in the literature as the universal model for this kind of observations (Mejail et al, 2003). Frey et al (2011) showed that using Information theory measures (as divergences and entropies) combined with statistical inference is a powerful methodology. Recently, Nikooravesh (2018) developed estimation procedures for the quantile function by means of Shannon and Tsallis entropies. We propose boundary detectors which are competitive with respect to those based on the joint likelihood, which are computationally demanding, as discussed by Nascimento et al (2014). We propose and discuss two boundary detection schemes based on the Kullback-Leibler (KL) and Rényi divergences. Additionally, we also investigate their performance at the limit case when intensities are gamma distributed. Results provide evidence in favor of the detector based on the Rényi divergence between \mathcal{G}_I^0 models.

The paper unfolds as follows. Section 2 recalls the \mathcal{G}_I^0 model. In Sections 3 and 4, we discuss the divergence measures and boundary detection procedures, respectively. Section 5 shows results of a simulation study and an application to an actual SAR image. Finally, Section 6 concludes the paper.

2. MODEL FOR SPECKLED DATA: THE \mathcal{G}_I^0 MODEL

The multiplicative model for the observation at position (i, j) of an intensity SAR image describes it as the outcome of the random variable $Z(i, j) = X(i, j)Y(i, j)$, where $X(i, j)$ and $Y(i, j)$ are independent random variables. The latter, which describes the speckle, follows

a unitary-mean gamma distribution with shape parameter $L \geq 1$; this parameter is known as “number of looks,” it is proportional to the signal-to-noise ratio, and it is often fixed for the whole image. The unobserved quantity of interest, $X(i, j)$, is called “backscatter.” The backscatter is positive, and contains all the relevant information about the target.

Assuming a Reciprocal Gamma law for the backscatter (Frery et al, 1997), we obtain that the density of Z is expressed by

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

where $\alpha < 0$ is the texture, and $\gamma > 0$ is the brightness. We denote this situation as $Z \sim \mathcal{G}^0(\alpha, \gamma, L)$, with $\Theta = \mathbb{R}_- \times \mathbb{R}_+ \times [1, \infty)$ the parameter space.

Frery et al (1997) proved the following result. Consider the sequence of random variables Z_1, Z_2, \dots in which $Z_i \sim \mathcal{G}_I^0(\alpha_i, \gamma_i, L)$. If $-\alpha_i, \gamma_i \rightarrow \infty$ such that $-\alpha_i/\gamma_i \rightarrow \beta_1$, then the following convergence in distribution holds:

$$Z_i \xrightarrow{\mathcal{D}} Z, \quad (1)$$

where Z follows a gamma distribution with mean β_1 and shape parameter L . In particular, if $L = 1$ then the convergence is towards an Exponential law.

Fig. 1 shows four single-look unitary-mean \mathcal{G}_I^0 densities with varying roughness: $\alpha \in \{-\infty, -10, -3, -1.5\}$. Since the \mathcal{G}_I^0 distribution is the Exponential law in the limit above, we plot this density in black to serve as a reference. The densities in linear scale might look like Exponential, but they are not, as revealed in the semi-logarithmic scale: \mathcal{G}_I^0 densities have heavier tails. The larger the texture parameter is, the heavier the tail is.

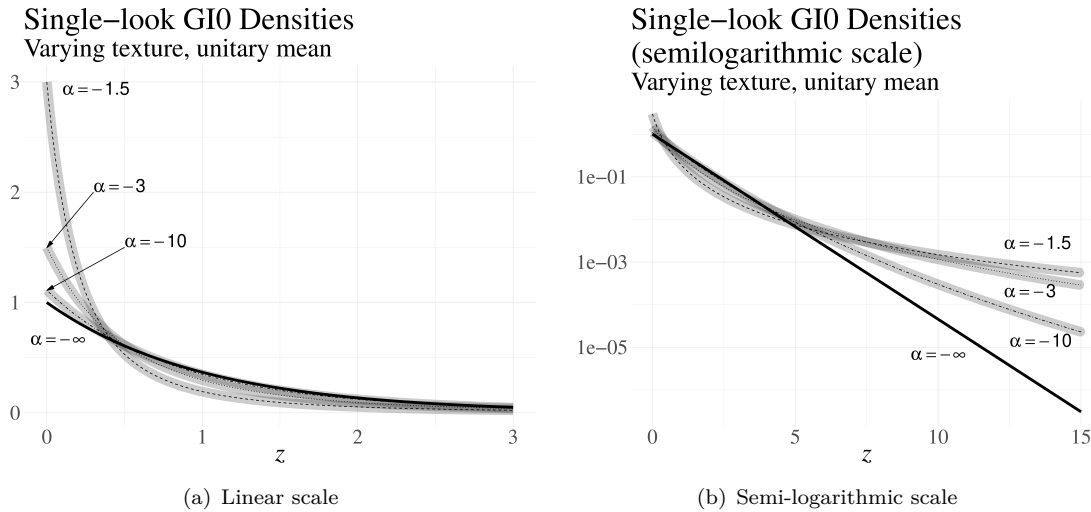
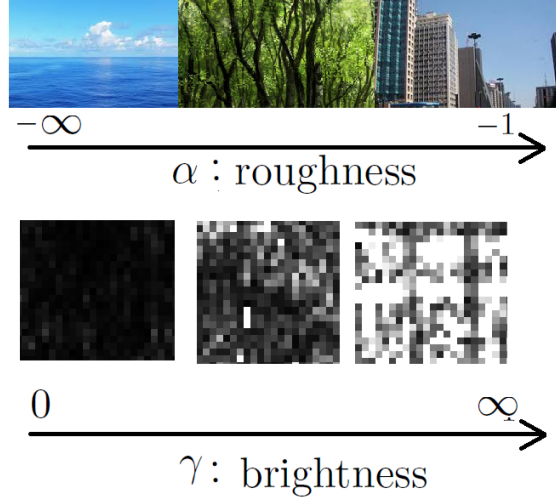


Figure 1. Single-look unitary-mean \mathcal{G}_I^0 densities with varying roughness.

A remarkable feature of this distribution is that it describes well extremely textured areas (urban centers), textured regions (forests), and areas with fully developed speckle and, thus, textureless (bare soil and crops, for instance). Figure 2 illustrates how α and γ can be interpreted.

Due to its desirable asymptotic properties (unbiasedness, normality, and efficiency), we use the maximum likelihood (ML) estimator for obtaining the parameters α and γ from data. Let $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ be a random sample drawn from $Z \sim \mathcal{G}_I^0(\alpha, \gamma, L)$. The

Figure 2. Properties of the \mathcal{G}_I^0 parameters.

likelihood function of the observed sample $\mathbf{z} = (z_1, z_2, \dots, z_n)$ is expressed as

$$\mathcal{L}(\alpha, \gamma; \mathbf{z}) = \left(\frac{L^L \Gamma(L - \alpha)}{\gamma^\alpha \Gamma(-\alpha) \Gamma(L)} \right)^n \prod_{i=1}^n z_i^{L-1} (\gamma + L z_i)^{\alpha-L}.$$

Assuming L fixed, the ML estimates for α and γ , say $\hat{\alpha}$ and $\hat{\gamma}$, respectively, are the solution of the following system of nonlinear equations:

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

where ψ^0 is the digamma function. This nonlinear system does not have a closed-form solution, then, we rely on numerical optimization methods.

3. CONTRAST BASED ON INFORMATION AND DIVERGENCE

The KL divergence (or relative entropy) is a well-known way of comparing two distributions. Divergence measures are submitted to a systematic and comprehensive treatment and, as a result, [Salicrú et al \(1994\)](#) proposed the class of (h, ϕ) -divergences.

Let Z_1 and Z_2 be two random variables equipped with densities f_{Z_1} and f_{Z_2} , respectively, with common support $I \subseteq \mathbb{R}$. The (h, ϕ) -divergence between Z_1 and Z_2 is defined by

$$D_\phi^h(Z_1 \| Z_2) = h \left(\int_I \phi \left(\frac{f_{Z_1}(z; \boldsymbol{\theta}_1)}{f_{Z_2}(z; \boldsymbol{\theta}_2)} \right) f_{Z_2}(z; \boldsymbol{\theta}_2) dz \right),$$

where $\phi: (0, \infty) \rightarrow [0, \infty)$ is a convex function, $h: (0, \infty) \rightarrow [0, \infty)$ is a strictly increasing function with $h(0) = 0$, and indeterminate forms are assigned value zero. Table 1 shows three choices of h and ϕ functions, and the resulting divergences.

Table 1. ϕ and h functions, and related divergences.

(h, ϕ) -distance/notation	$h(y)$	$\phi(x)$
Kullback-Leibler/KL	$y/2$	$(x-1) \log(x)$
Rényi (order $0 < \beta < 1$)/RD: β	$\frac{1}{\beta-1} \log((\beta-1)y+1), 0 \leq y < \frac{1}{1-\beta}$	$\frac{x^{1-\beta} + x^\beta - \beta(x-1) - 2}{2(\beta-1)}$
Bhattacharyya/BA	$-\log(-y+1), 0 \leq y < 1$	$-\sqrt{x} + \frac{x+1}{2}$

In particular, consider Z_1 and Z_2 be random variables following the \mathcal{G}_I^0 model and indexed by parameters $\boldsymbol{\theta}_1 = [\alpha_1, \gamma_1, L_1]^\top$ and $\boldsymbol{\theta}_2 = [\alpha_2, \gamma_2, L_2]^\top$, respectively. The KL divergence between Z_1 and Z_2 can be computed setting $A(\alpha, \gamma, L) \equiv L^\alpha \Gamma(L - \alpha) [\gamma^\alpha \Gamma(-\alpha) \Gamma(L)]^{-1}$. With this

$$\begin{aligned}
D_{\text{KL}}^{\mathcal{G}_I^0}(Z_1 \| Z_2) &= \int_0^\infty f_{Z_1}(z) \log \left(\frac{f_{Z_1}(z)}{f_{Z_2}(z)} \right) dz \\
&= \int_0^\infty A(\alpha_1, \gamma_1, L_1) z^{(L_1-1)} (\gamma_1 + L_1 z)^{\alpha_1-L_1} \left[\log \left(\frac{A(\alpha_1, \gamma_1, L_1)}{A(\alpha_2, \gamma_2, L_2)} \right) \right. \\
&\quad \left. + \log \left(\frac{z^{L_1-1} (\gamma_1 + L_1 z)^{\alpha_1-L_1}}{z^{L_2-1} (\gamma_2 + L_2 z)^{\alpha_2-L_2}} \right) \right] dz \\
&= \log \left(\frac{A(\alpha_1, \gamma_1, L_1)}{A(\alpha_2, \gamma_2, L_2)} \right) + (L_1 + L_2 - 2) \mathbb{E}[\log(Z_1)] + (\alpha_1 - L_1) \mathbb{E}[\log(\gamma_1 + L_1 Z_1)] \\
&\quad - (\alpha_2 - L_2) \mathbb{E}[\log(\gamma_2 + L_2 Z_1)],
\end{aligned}$$

where \mathbb{E} denotes the expected value, $\mathbb{E}[\log(\gamma_1 + L_1 Z_1)] = \log(\gamma_1) + \psi(L - \alpha_1) - \psi(-\alpha_1)$, $\mathbb{E}[\log(Z_1)] = \log(\gamma_1) - \psi(-\alpha_1) + \psi(L) - \Gamma(L)^{-1} [\log(L) + 1] - 1$ and $\mathbb{E}[\log(\gamma_2 + L_2 Z_1)]$ is a quantifly which can be defined in terms of the following integral

$$\begin{aligned}
&\int_0^\infty \log(\gamma_2 + L_2 z) z^{L_1-1} (\gamma_1 + L_1 z)^{\alpha_1-L_1} dz = -\frac{L_2^{-\alpha_1} [L_1 (\gamma_2 L_1 - \gamma_1 L_2)]^{-L_1}}{\alpha_1 (\alpha_1^2 - 1) \gamma_2 L_1 \Gamma(L_1 - \alpha_1 + 1)} \\
&\times \left\{ -\frac{(L_1 - \alpha_1) (\gamma_2 L_1 - \gamma_1 L_2)^{L_1} \Gamma(L_1 + 1) (\gamma_1 L_2)^{\alpha_1+1} {}_3F_2 \left(1, 1, L_1 + 1; 2, \alpha_1 + 2; \frac{L_2 \gamma_1}{L_1 \gamma_2} \right)}{\Gamma(\alpha_1 - 1)} \right. \\
&\times \pi \csc(\pi \alpha_1) + (\alpha_1 + 1) (\alpha_1 - L_1) (\gamma_2 L_1 - \gamma_1 L_2)^{L_1} \left[\gamma_2 \Gamma(2 - \alpha_1) \log(\gamma_2) \Gamma(L_1 + 1) (\gamma_1 L_2)^{\alpha_1} \right. \\
&+ \pi \gamma_1 L_2 \csc(\pi \alpha_1) (\gamma_2 L_1)^{\alpha_1} \Gamma(L_1 - \alpha_1 + 1) {}_2F_1 \left(1 - \alpha_1, L_1 - \alpha_1 + 1; 2 - \alpha_1; \frac{L_2 \gamma_1}{L_1 \gamma_2} \right) \left. \right] \\
&\left. - \pi (\alpha_1^2 - 1) \csc(\pi \alpha_1) (\gamma_2 L_1)^{L_1+1} (\gamma_2 L_1 - \gamma_1 L_2)^{\alpha_1} \Gamma(L_1 - \alpha_1 + 1) \right\},
\end{aligned}$$

where ${}_pF_q$ is the generalized hypergeometric function (Gradshteyn and Ryzhik, 1980, Sec. 9.18). Under the conditions of Equation (1), Equation (2) collapsed to

$$\begin{aligned}
D_{\text{KL}}^\Gamma(Z_1 \| Z_2) &= L_1 \log(L_1/\beta_1) - L_2 \log(L_2/\beta_2) + \log(\Gamma(L_2)) - \log(\Gamma(L_1)) \\
&\quad + L_1 \left(\frac{L_2/\beta_2}{L_1/\beta_1} - 1 \right) + [\psi(L_1) - \log(L_1/\beta_1)](L_1 - L_2).
\end{aligned}$$

One can note that both $D_{\text{KL}}^{\mathcal{G}_I^0}(\cdot \| \cdot)$ and $D_{\text{KL}}^\Gamma(\cdot \| \cdot)$ are non-symmetric measurers. A simple solution for addressing the symmetry problem is the definition of a new measure d_ϕ^h expressed

by

$$d_\phi^h(Z_1, Z_2) = \frac{D_\phi^h(Z_1 \| Z_2) + D_\phi^h(Z_2 \| Z_1)}{2}. \quad (2)$$

In this paper, we work with the KL distances stated as

$$d_{\text{KL}}^{\mathcal{G}_I^0}(Z_1, Z_2) = \frac{D_{\text{KL}}^{\mathcal{G}_I^0}(Z_1 \| Z_2) + D_{\text{KL}}^{\mathcal{G}_I^0}(Z_2 \| Z_1)}{2}$$

and

$$d_{\text{KL}}^\Gamma(Z_1, Z_2) = \frac{D_{\text{KL}}^\Gamma(Z_1 \| Z_2) + D_{\text{KL}}^\Gamma(Z_2 \| Z_1)}{2}.$$

We also consider the Rényi distance (of order $0 < \beta < 1$) between \mathcal{G}_I^0 distributions given by

$$\begin{aligned} (\beta - 1) d_{\text{RE};\beta}^{\mathcal{G}_I^0}(Z_1, Z_2) &= \log \left(\int_0^\infty \frac{1}{2} [f_{Z_1}^\beta(z) f_{Z_2}^{1-\beta}(z) + f_{Z_2}^\beta(z) f_{Z_1}^{1-\beta}(z)] dz \right) \\ &= -\log(2) + \log \left[\int_0^\infty x^{\beta L_1 + (1-\beta)L_2 - 1} (\gamma_1 + L_1 x)^{\beta(\alpha_1 - L_1)} \right. \\ &\quad \times (\gamma_2 + L_2 x)^{(1-\beta)(\alpha_2 - L_2)} dx \times A(\alpha_1, \gamma_1, L_1)^\beta A(\alpha_2, \gamma_2, L_2)^{1-\beta} \\ &\quad \left. + \int_0^\infty x^{\beta L_2 + (1-\beta)L_1 - 1} \times (\gamma_2 + L_2 x)^{\beta(\alpha_2 - L_2)} (\gamma_1 + L_1 x)^{(1-\beta)(\alpha_1 - L_1)} dx \right. \\ &\quad \left. \times A(\alpha_2, \gamma_2, L_2)^\beta A(\alpha_1, \gamma_1, L_1)^{1-\beta} \right]. \end{aligned}$$

The expressions of the above integrations are suppressed for simplicity. Under the conditions of Equation (1), we have that

$$\begin{aligned} 2 \exp\{(\beta - 1) d_{\text{RE};\beta}^\Gamma(Z_1, Z_2)\} &= \left(\frac{\Gamma(L_2)}{(L_2/\beta_2)^{L_2}} \frac{(L_1/\beta_1)^{L_1}}{\Gamma(L_1)} \right)^{\beta-1} \\ &\quad \times \Gamma(\underbrace{\beta L_1 + (1-\beta)L_2}_{L_{01}}) \frac{[\beta(L_1/\beta_1) + (1-\beta)(L_2/\beta_2)]^{-L_{01}}}{(1/(\beta_1/L_1))^{L_1} \Gamma(L_1)} \\ &\quad + \left(\frac{\Gamma(L_1)}{(L_1/\beta_1)^{L_1}} \frac{(L_2/\beta_2)^{L_2}}{\Gamma(L_2)} \right)^{\beta-1} \Gamma(\underbrace{\beta L_2 + (1-\beta)L_1}_{L_{02}}) \\ &\quad \times \frac{[\beta(L_2/\beta_2) + (1-\beta)(L_1/\beta_1)]^{-L_{02}}}{(1/(\beta_2/L_2))^{L_2} \Gamma(L_2)}. \end{aligned}$$

In particular, one obtains the Bhattacharyya distance when $\beta = 1/2$ and the final expression is multiplied by $1/2$

$$\begin{aligned} d_{\text{BA}}^\Gamma(Z_1, Z_2) &= -\log \left(\int_0^\infty \sqrt{f_{Z_1}(z; L_1, L_1 \beta_1) f_{Z_2}(z; L_2, L_2 \beta_2)} dz \right) \\ &= -\frac{L_1 \log(L_1/\beta_1) + L_2 \log(L_2/\beta_2)}{2} + \left(\frac{L_1 + L_2}{2} \right) \log \left(\frac{(L_1/\beta_1) + (L_2/\beta_2)}{2} \right) \\ &\quad - \log \left(\Gamma \left(\frac{L_1 + L_2}{2} \right) \right) + \left[\frac{\log \Gamma(L_1) + \log \Gamma(L_2)}{2} \right]. \end{aligned}$$

The Hellinger (H) distance between gamma distributions can be derived from $d_{\text{BA}}^\Gamma(Z_1, Z_2)$ as

$$d_{\text{H}}^\Gamma(Z_1, Z_2) = 1 - e^{-d_{\text{BA}}^\Gamma(Z_1, Z_2)}.$$

We now employ these measures to propose new boundary detection tools. Although the previous contrast discussion considers that the number of looks are different and unknown, called “equivalent number of looks”, this parameter can be assumed common and known in the whole image. From now on, we assume it is a known constant for the \mathcal{G}_I^0 law and an estimable unknown constant for the Γ distribution as approached by [Anfinssen et al \(2009\)](#). Under this setting, both models have the same parametric space dimension being, thus, comparable.

4. BOUNDARY DETECTORS

An edge detector seeks a point on a strip of data where the statistical properties change. The detection procedures used in this paper work in three stages: (i) identifying the centroid of the candidate area (in automatic, semiautomatic, or manual manner), (ii) detecting transition points which belong to the edge, and (iii) defining the contour using a imputation method among the transition points, such as B-Splines ([Gambini et al, 2006](#)). We focus our analysis on stages (ii) and (iii).

Assume that an initial region \mathcal{R} with centroid C is available. Rays are traced from C to points outside \mathcal{R} . They are of the form $\mathbf{s}^{(i)} = \overline{CP_i}$, where the angle between rays is $\angle(\mathbf{s}^{(i)}, \mathbf{s}^{(i+1)})$, for $i = 1, 2, \dots, S$, being S the number of rays. Finally, the data are collected in thin strips around these rays.

We assume that the data follow a \mathcal{G}_I^0 distribution, and that there are two populations: one inside the edge with $j^{(i)}$ observations, and another outside the edge with $N^{(i)} - j^{(i)}$ observations. We can then model the $N^{(i)}$ observations around segment $\mathbf{s}^{(i)}$, $1 \leq i \leq S$ as

$$\begin{cases} Z_k^{(i)} \sim \mathcal{G}_I^0(\alpha_A^{(i)}, \gamma_A^{(i)}, L), & \text{for } k = 1, \dots, j^{(i)}, \\ Z_k^{(i)} \sim \mathcal{G}_I^0(\alpha_B^{(i)}, \gamma_B^{(i)}, L), & \text{for } k = j^{(i)} + 1, \dots, N^{(i)}. \end{cases} \quad (3)$$

In the limit case, that is, under the conditions of Equation (1), Equation (3) becomes

$$\begin{cases} Z_k^{(i)} \sim \Gamma(L_A^{(i)}, L_A^{(i)}/\beta_A^{(i)}), & \text{for } k = 1, \dots, j^{(i)}, \\ Z_k^{(i)} \sim \Gamma(L_B^{(i)}, L_B^{(i)}/\beta_B^{(i)}), & \text{for } k = j^{(i)} + 1, \dots, N^{(i)}. \end{cases} \quad (4)$$

Note that Equation (3) collapses in Equation (4) if $L_A^{(i)} = L_B^{(i)} = L$, $-\alpha_A^{(i)}/\gamma_A^{(i)} \rightarrow \beta_A^{(i)}$ and $-\alpha_B^{(i)}/\gamma_B^{(i)} \rightarrow \beta_B^{(i)}$, but these laws are not nested and, therefore, are competitive.

The main idea is to find the edge $j^{(i)}$ th on the segment $\mathbf{s}^{(i)}$ as the point that provides the best configuration according with respect to a decision rule. In the following we present three different decision rules omitting, for the sake of brevity, the index (i) since only one strip is considered at each epoch.

The log-likelihood for the configuration stated in Equation (3) (or (4)) is given by

$$\ell(j) = \log(L(j)) = \sum_{k=1}^j \log(f_{Z_1}(z_k; \boldsymbol{\theta}_A)) + \sum_{k=j+1}^N \log(f_{Z_2}(z_k; \boldsymbol{\theta}_B)),$$

where $\boldsymbol{\theta}_A \in \{[\alpha_A, \gamma_A], [L_A, \beta_A]\}$ and $\boldsymbol{\theta}_B \in \{[\alpha_B, \gamma_B], [L_B, \beta_B]\}$.

Gambini et al (2006) showed that an efficient estimator, \hat{j}_{ML} , for the index on the segment that corresponds to the transition point is stated as $\hat{j}_{\text{ML}} = \arg \max_j \ell(j)$. However, this procedure is computationally demanding as it needs to evaluate two likelihood functions at each search step.

As discussed in Nascimento et al (2010), the distances derived in Section 3 can be scaled to be asymptotically distributed as chi-square statistics:

$$S_{\mathcal{D}}(\widehat{\boldsymbol{\theta}}_1(j), \widehat{\boldsymbol{\theta}}_2(N-j)) = \frac{2j(N-j)v_{\mathcal{D}}}{N} d_{\mathcal{D}}(\widehat{\boldsymbol{\theta}}_1(j), \widehat{\boldsymbol{\theta}}_2(N-j)),$$

where $v_{\mathcal{D}} = 1, \beta^{-1}, 4$, and 4 for $\mathcal{D} = \text{KL}, \text{RD}:\beta, \text{BA}$, and H , respectively, and $\widehat{\boldsymbol{\theta}}_1(j) = [\widehat{\alpha}_A(j), \widehat{\gamma}_A(j)]$ and $\widehat{\boldsymbol{\theta}}_2(N-j) = [\widehat{\alpha}_B(N-j), \widehat{\gamma}_B(N-j)]$ are the maximum likelihood estimators for $\boldsymbol{\theta}_1 = (\alpha_A, \gamma_A)$ and $\boldsymbol{\theta}_2 = (\alpha_B, \gamma_B)$ using random samples of sizes j and $N-j$, respectively. Under mild conditions, $S_{\mathcal{D}}(\widehat{\boldsymbol{\theta}}_1(j), \widehat{\boldsymbol{\theta}}_2(N-j))$ is asymptotically distributed as a χ^2_2 random variable under the null hypothesis $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$.

Thus, we propose novel detectors for finding edges on SAR intensities by seeking for the point that maximizes the test statistics between the two models, that is,

$$\hat{j}_{\mathcal{D}} = \arg \max_j S_{\mathcal{D}}(\widehat{\boldsymbol{\theta}}_1(j), \widehat{\boldsymbol{\theta}}_2(N-j)) = \arg \max_j S_{\mathcal{D}}(j),$$

where $\mathcal{D} = \{\text{KL}; \text{BA}; \text{H}; \text{RD}:\beta\}$.

5. NUMERICAL RESULTS

In the simulations, we study the performance of the two edge detectors here proposed. We use simulated data from two models, namely gamma and \mathcal{G}_I^0 and edge detectors based on the gamma and \mathcal{G}_I^0 distributions. With this, we verify the robustness of the detectors when fed with data that do not belong to the model they were originally devised. We utilize the absolute value of the difference between the mean detection and actual edge position as performance criterion, given by

$$D = |\bar{B} - 100|,$$

where \bar{B} is the sample mean of detected edge and 100 is the true edge position. The smaller this measure is, the better the performance is.

We performed a Monte Carlo simulation study with: i) 1000 replications for each situation; ii) \mathcal{G}_I^0 and Γ distributed data; iii) in each replicate we simulate a strip of data of size 1×200 . The first half from one distribution, and the second from another distribution.

The first set of experiments used edge detectors based on the gamma distribution, in which we set $\beta = 0.9$ in the Rényi distance. The observations are samples Z_1, Z_2, \dots, Z_{100} from $\Gamma(\nu_0, \nu_0/\beta_0)$, and $Z_{101}, Z_{102}, \dots, Z_{200}$ from $\Gamma(\nu_1, \nu_1/\beta_1)$, with $\beta_0 = \beta_1 = 1$ (unitary mean), $\nu_0 = 1$ and $\nu_1 = 4, 6, 8$.

The second set of experiments used edge detectors based on the \mathcal{G}_I^0 law and samples Z_1, Z_2, \dots, Z_{100} from $\mathcal{G}_I^0(-\nu_0, \nu_0, 4)$, and $Z_{101}, Z_{102}, \dots, Z_{200}$ from $\mathcal{G}_I^0(-\nu_1, \nu_1, 4)$, with $\nu_0 = 1.5, 3$ and $\nu_1 = 3, 5$. These parameters provide a small but representative set of values.

Table 2 shows the performance of the Gamma-based edge detectors. The best results (smallest errors) are highlighted in gray. In this table, finding an edge between Gamma samples seems a difficult task for Gamma-based detectors. The Hellinger-based detector performs best in call cases, but the mean errors are of the order of four pixels. The Rényi-based detector performs worst.

Table 2. Performance measures for gamma distances in the indicated model.

Distribution and parameters	gamma-based detectors				
	LR	KL	BA	H	RD, β
Gamma, $\nu_0 = 1, \nu_1 = 4$	26.51	6.41	11.35	5.05	60.50
Gamma, $\nu_0 = 1, \nu_1 = 6$	28.55	4.71	10.18	4.41	66.93
Gamma, $\nu_0 = 1, \nu_1 = 8$	29.64	4.68	10.88	3.96	68.94
$\mathcal{G}_0^I, \nu_0 = 1.5, \nu_1 = 3$	27.57	12.26	13.73	7.01	4.93
$\mathcal{G}_0^I, \nu_0 = 1.5, \nu_1 = 5$	18.19	1.25	20.80	0.14	1.70
$\mathcal{G}_0^I, \nu_0 = 3, \nu_1 = 5$	26.75	1.56	19.98	8.65	6.44

Table 3 shows the performance of the \mathcal{G}_I^0 -based edge detectors. In this table, overall, the two edge detectors based on the \mathcal{G}_I^0 distribution have similar performance. The Rényi-based detector produces consistently better results than the KL-based one when the data follow \mathcal{G}_I^0 laws. Although the latter detector is better in two out of three cases of Gamma-distributed samples, the differences are approximately 1 % and 7 %.

Table 3. Performance measures for \mathcal{G}_0^I distances.

Distribution and parameters	\mathcal{G}_0^I -based detectors	
	KL	RD, β
Gamma, $\nu_0 = 1, \nu_1 = 4$	6.30	6.68
Gamma, $\nu_0 = 1, \nu_1 = 6$	9.84	9.94
Gamma, $\nu_0 = 1, \nu_1 = 8$	10.70	10.20
$\mathcal{G}_0^I, \nu_0 = 1.5, \nu_1 = 3$	3.23	1.27
$\mathcal{G}_0^I, \nu_0 = 1.5, \nu_1 = 5$	5.12	1.33
$\mathcal{G}_0^I, \nu_0 = 3, \nu_1 = 5$	3.57	1.05

The results presented in Tables 2 and 3 led us to conclude that the safest and most versatile option for edge detection is the \mathcal{G}_I^0 -based detector that uses the Rényi distance with $\beta = 0.9$. We now are in position of submitting the detectors to real data. We present an application to an actual SAR image to assess the proposed detectors in practice.

Figure 3(a) displays a SAR image of crops in Foulum (Denmark) from the HH (horizontal-horizontal) polarization channel. This picture has been obtained by an EMISAR sensor with four nominal looks. Figure 3(b) shows the reference map, and Figure 3(c) identifies classes with shades of gray. According to the discussion about this image Foulum in Ferreira and Nascimento (2020), there is a centroid between the wheat and rapeseed areas. We use it to cast the rays on which the proposed detectors work. Figures 3(d) and 3(e) exhibit the result of detecting the edges that separate wheat from rapeseed.

The following analysis is made by visual inspection on the edges reconstructed from the estimated transition points and fourth-degree B-splines curves, as in Nascimento et al (2014).

The performance of \mathcal{G}_I^0 -based detectors is consistently better than those obtained from the Γ law. They provide the same good estimate of the edge. This result is in agreement with the simulation study once the \mathcal{G}_I^0 distribution is a better alternative to describe different SAR clutter. The Rényi distance furnishes the best detection for the Gamma-based detectors.

Simulation and real experiments were made in the R programming environment (Wickham, 2019); functions `integrate` and `maxLik(. , method=BFGS)` were used for numerical integration and obtaining ML estimates (equipped with moments method ones as initial point), respectively. All studies were performed in a Intel(R) Core(TM) i5-5200U processor at 2.20 GHz.

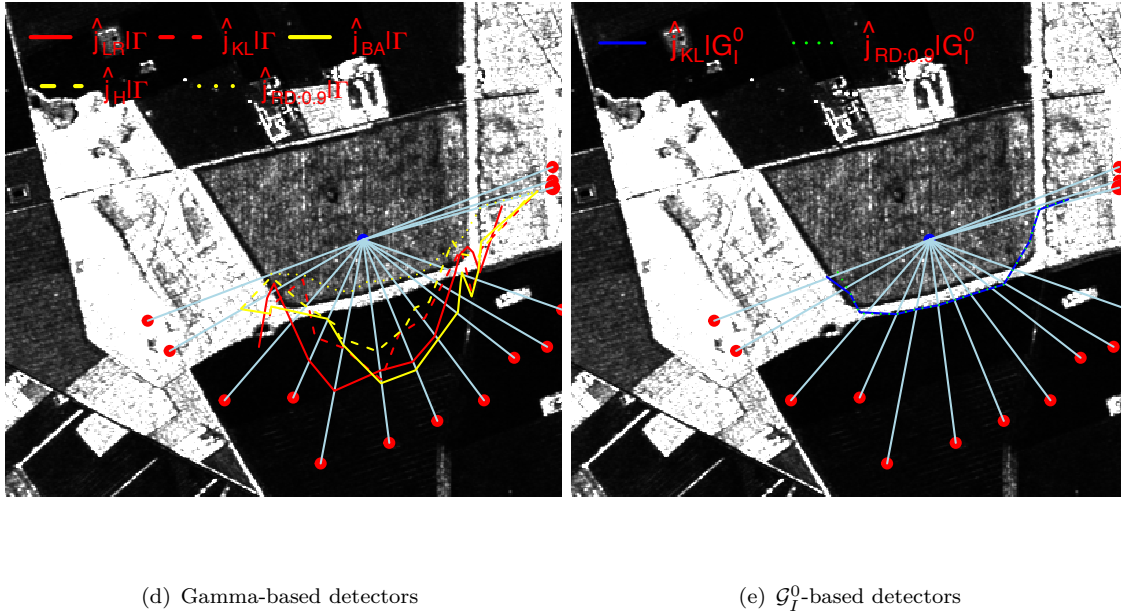
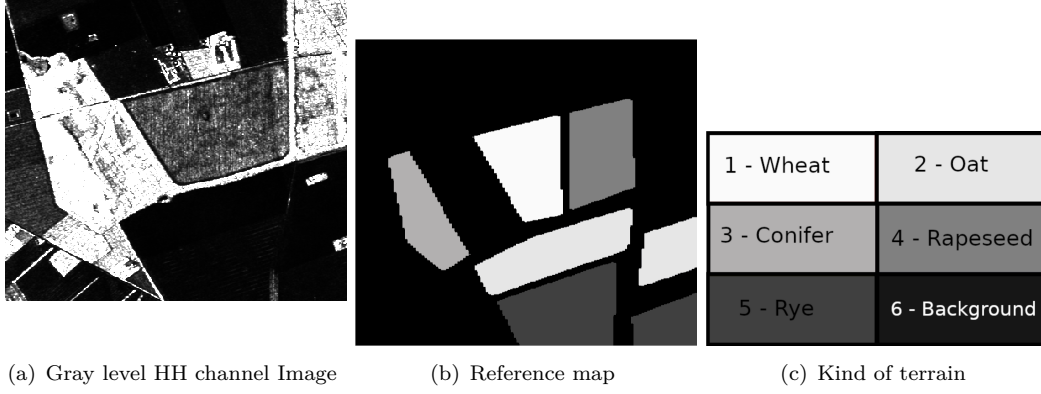


Figure 3. Application of edge detectors based on stochastic distances on an actual SAR image.

6. CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

In this paper, we proposed distance-based boundary detectors for synthetic aperture radar data modeled by the Γ and \mathcal{G}_I^0 laws. These proposals have wide applicability in practice, like the Monitoring of oil spill (Fan et al, 2015) and deforestation areas in forests (Bouvet et al, 2018). We quantified their performances with both simulated and actual data. Results provided evidence that detectors based on Kullbak-Leibler and Rényi distances for \mathcal{G}_I^0 models outperform ones based on distances between Γ limit cases and on the joint Γ likelihood, which have been employed in the synthetic aperture radar literature.

This paper has addressed only measures induced from the Γ and \mathcal{G}_I^0 distributions and approached the univariate aspect, that consist in two of its limitations.

In future works, we use the general distributions like complete \mathcal{G} (Frery et al, 1997) and KummerU (Deng et al, 2017).

REFERENCES

- Anfinssen, S.N., Doulgeris, A.P., and Eltoft, T., 2009. Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 47, 3795–3809.
- Bouvet, A., Mermoz, S., Ballère, M., Koleček, T., and Toan, T.L., 2018. Use of the SAR Shadowing Effect for Deforestation Detection with Sentinel-1 Time Series. *Remote Sensing*, 10, 1–20.
- Deng, X., López-Martínez, C., Chen, J., and Han, P., 2017. Statistical modeling of polarimetric SAR data: A survey and challenges. *Remote Sensing*, 9, 1–34.
- Fan, J., Zhang, F., Zhao, D., and Wang, J., 2015. Oil Spill Monitoring Based on SAR Remote Sensing Imagery. *Aquatic Procedia*, 3, 112–118.
- Ferreira, J.A. and Nascimento, A.D.C., 2020. Shannon entropy for the \mathcal{G}_I^0 model: A new segmentation approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 2547–2553.
- Frery, A.C., Muller, H.J., Yanasse, C.C.F., and Sant’Anna, S.J.S., 1997. A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*, 35, 648–659.
- Frery, A.C., Nascimento, A.D.C., and Cintra, R.J., 2011. Information theory and image understanding: An application to polarimetric SAR imagery. *Chilean Journal of Statistics*, 2, 81–100.
- Gambini, J., Mejail, M., Jacobo-Berlles, J., and Frery, A.C., 2006. Feature extraction in speckled imagery using dynamic B-spline deformable contours under the G0 model. *International Journal of Remote Sensing*, 27, 5037–5059.
- Gambini, J., Mejail, M., Jacobo-Berlles, J., and Frery, A.C., 2008. Accuracy of edge detection methods with local information in speckled imagery. *Statistics and Computing*, 18, 15–26.
- Giron, E., Frery, A.C., and Cribari-Neto, F., 2012. Nonparametric edge detection in speckled imagery. *Mathematics and Computers in Simulation*, 82, 2182–2198.
- Gradshteyn, I.S. and Ryzhik, I.M., 1980. *Tables of Integrals, Series and Products*. Academic Press, New York.
- Kong, J., 1990. *Electromagnetic Wave Theory*. Wiley, New York.
- Mejail, M.E., Jacobo-Berlles, J., Frery, A.C., and Bustos, O.H., 2003. Classification of SAR images using a general and tractable multiplicative model. *International Journal of Remote Sensing*, 24, 3565–3582.
- Nascimento, A.D.C., Cintra, R.J., and Frery, A.C., 2010. Hypothesis testing in speckled data with stochastic distances. *IEEE Transactions on Geoscience and Remote Sensing*, 48, 373–385.
- Nascimento, A.D.C., Horta, M.M., and Frery, A.C., Cintra, R.J., 2014. Comparing edge detection methods based on stochastic entropies and distances for PolSAR imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7, 648–663.
- Nikooravesh, Z., 2018. Estimation of the probability function under special moments conditions using the maximum Shannon and Tsallis entropies. *Chilean Journal of Statistics*, 9, 55–64.
- Salicrú, M., Menéndez, M.L., Pardo, L., and Morales, D., 1994. On the applications of divergence type measures in testing statistical hypothesis. *Journal of Multivariate Analysis*, 51, 372–391.
- Sant’Anna, S.J.S., Lacava, J.C.S., and Fernandes, D., 2008. From Maxwell’s equations to polarimetric SAR images: A simulation approach. *Sensors*, 8, 7380–7409.

- Ward, K.D., 1981. Compound representation of high resolution sea clutter. *Electronic Letters*, 17, 561-563.
- Ward, K.D., Watts, S., and Tough, R.J., 2006. Sea clutter: scattering, the K distribution and radar performance. *IET*, 20.
- Wei, Q.R., Feng, D.Z. 2015. An efficient SAR edge detector with a lower false positive rate. *International Journal of Remote Sensing*, 36, 3773-3797.
- Wickham, H. 2019. *Advanced R*. CRC Press, Boca Raton, FL, US.
- Yue, D.X., Xu, F., Frery, A.C., and Jin, Y.Q. 2020. A generalized Gaussian coherent scatterer model for correlated SAR texture. *IEEE Transactions on Geoscience and Remote Sensing*, 58, 2947-2964.
- Yue, D.X., Xu, F., Frery, A.C., and Jin, Y.Q., 2021. SAR image statistical modeling Part I: Single-pixel statistical models. *IEEE Geoscience and Remote Sensing Magazine*, 9, 82-114.

INFORMATION FOR AUTHORS

The editorial board of the Chilean Journal of Statistics (ChJS) is seeking papers, which will be refereed. We encourage the authors to submit a PDF electronic version of the manuscript in a free format to the Editors-in-Chief of the ChJS (E-mail: chilean.journal.of.statistics@gmail.com). Submitted manuscripts must be written in English and contain the name and affiliation of each author followed by a leading abstract and keywords. The authors must include a "cover letter" presenting their manuscript and mentioning: "We confirm that this manuscript has been read and approved by all named authors. In addition, we declare that the manuscript is original and it is not being published or submitted for publication elsewhere".

PREPARATION OF ACCEPTED MANUSCRIPTS

Manuscripts accepted in the ChJS must be prepared in Latex using the ChJS format. The Latex template and ChJS class files for preparation of accepted manuscripts are available at <http://soche.cl/chjs/files/ChJS.zip>. Such as its submitted version, manuscripts accepted in the ChJS must be written in English and contain the name and affiliation of each author, followed by a leading abstract and keywords, but now mathematics subject classification (primary and secondary) are required. AMS classification is available at <http://www.ams.org/mathscinet/msc/>. Sections must be numbered 1, 2, etc., where Section 1 is the introduction part. References must be collected at the end of the manuscript in alphabetical order as in the following examples:

Arellano-Valle, R., 1994. Elliptical Distributions: Properties, Inference and Applications in Regression Models. Unpublished Ph.D. Thesis. Department of Statistics, University of São Paulo, Brazil.

Cook, R.D., 1997. Local influence. In Kotz, S., Read, C.B., and Banks, D.L. (Eds.), Encyclopedia of Statistical Sciences, Vol. 1., Wiley, New York, pp. 380-385.

Rukhin, A.L., 2009. Identities for negative moments of quadratic forms in normal variables. Statistics and Probability Letters, 79, 1004-1007.

Stein, M.L., 1999. Statistical Interpolation of Spatial Data: Some Theory for Kriging. Springer, New York.

Tsay, R.S., Peña, D., and Pankratz, A.E., 2000. Outliers in multivariate time series. Biometrika, 87, 789-804.

References in the text must be given by the author's name and year of publication, e.g., Gelfand and Smith (1990). In the case of more than two authors, the citation must be written as Tsay et al. (2000).

COPYRIGHT

Authors who publish their articles in the ChJS automatically transfer their copyright to the Chilean Statistical Society. This enables full copyright protection and wide dissemination of the articles and the journal in any format. The ChJS grants permission to use figures, tables and brief extracts from its collection of articles in scientific and educational works, in which case the source that provides these issues (Chilean Journal of Statistics) must be clearly acknowledged.