Identifying Departures from the Fully Developed Speckle Hypothesis in Intensity SAR Data with Non-Parametric Estimation of the Entropy

Janeth Alpala Abraão D. C. Nascimento **Alejandro C. Frery**

Universidade Federal de Pernambuco

Victoria University of Wellington
Te Herenga Waka

MIGARS 2024
International Conference on Machine Intelligence for Geoanalytics and Remote Sensing
8–10 April 2024
Wellington • New Zealand

Outline

Context and Problem Statement

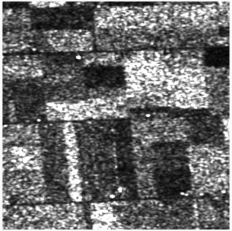
Context

Fully Developed Speckle (FDS)

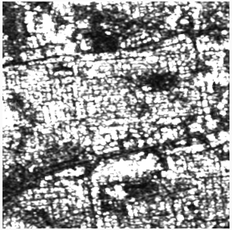
Homogeneous areas in SAR images are characterized by the lack of dominant scatterers, and the surface can be considered stationary. This is the case of the **fully developed hypothesis for the speckle**.

Optimal model

The $\mathcal{G}^0$ distribution is a suitable model for SAR intensity data. It includes the **Gamma law** as a limit case that results in the presence of fully developed speckle.



(a) FDS (agricultural area)



(b) Heterogeneous clutter (urban area)

Problem and Proposal

Selecting the optimal statistical model to characterize speckle in SAR data presents challenges.

Problem and Proposal

Selecting the optimal statistical model to characterize speckle in SAR data presents challenges.

- Opting for the **Gamma law** with $\mathcal{G}^0$ data:
    - Loss of the information about the number of scatterers.

Problem and Proposal

Selecting the optimal statistical model to characterize speckle in SAR data presents challenges.

- Opting for the **Gamma law** with $\mathcal{G}^0$ data:
  - Loss of the information about the number of scatterers.
- **Applying** $\mathcal{G}^0$ under fully developed speckle:
  - Tricky maximum likelihood estimation: Increased bias, flat likelihood, and numerical optimization may not converge.

Problem and Proposal

Selecting the optimal statistical model to characterize speckle in SAR data presents challenges.

- Opting for the **Gamma law** with $\mathcal{G}^0$ data:
  - Loss of the information about the number of scatterers.
- **Applying** $\mathcal{G}^0$ under fully developed speckle:
  - Tricky maximum likelihood estimation: Increased bias, flat likelihood, and numerical optimization may not converge.

Our approach

We propose a test statistic to distinguish between fully developed speckle and heterogeneous clutter, based on the non-parametric entropy estimation.

Model Setup

Model Setup

Intensity SAR Data

We denote $Z \sim \Gamma_{\mathsf{SAR}}(L, \mu)$ and $Z \sim \mathcal{G}_I^0(\alpha, \gamma, L)$ to indicate that Z follows the distributions characterized by the respective probability density functions:

$$f_Z(z; L, \mu) = \frac{L^L}{\Gamma(L)\mu^L} z^{L-1} \exp\left\{-Lz/\mu\right\} \mathbb{1}_{\mathbb{R}_+}(z), \tag{1}$$

$$f_Z(z; \alpha, \gamma, L) = \frac{L^L \Gamma(L-\alpha)}{\gamma^\alpha \Gamma(-\alpha)\Gamma(L)} \cdot \frac{z^{L-1}}{(\gamma + Lz)^{L-\alpha}} \mathbb{1}_{\mathbb{R}_+}(z), \tag{2}$$

where $L \geqslant 1$ is the number of looks, $\Gamma(\cdot)$ is the gamma function, and $\mathbb{1}_A(z)$ is the indicator function of the set A. In (1), $\mu > 0$ is the mean; in (2) $\gamma > 0$ is the scale, $\alpha < -1$ measures the roughness.

Model Setup

New Parametrization of $\mathcal{G}_I^0$

We can parametrize (2) by the mean value:

$$\mu = -\frac{\gamma}{\alpha + 1}.$$

Model Setup

New Parametrization of $\mathcal{G}_I^0$

We can parametrize (2) by the mean value:

$$\mu = -\frac{\gamma}{\alpha+1}.$$

Thus, the probability density function that characterize the $G_I^0(\mu,\alpha,L)$ law is

$$f_Z(z;\mu,\alpha,L) = \frac{L^L\Gamma(L-\alpha)}{\big(-\mu(\alpha+1)\big)^{\alpha}\Gamma(-\alpha)\Gamma(L)}\cdot\frac{z^{L-1}}{\big(-\mu(\alpha+1)+Lz\big)^{L-\alpha}}\mathbb{1}_{\mathbb{R}_+}(z).$$

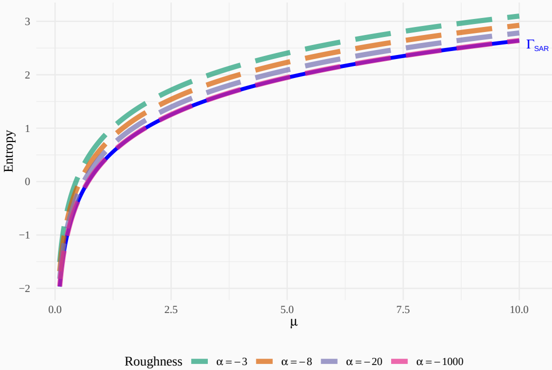
Non-parametric Entropy Estimation Approach

Shannon Entropy

$$H_{\Gamma_{\text{SAR}}}(L,\mu) = L - \ln L + \ln\Gamma(L) + (1-L)\psi^{(0)}(L) + \ln\mu,$$

$$\begin{aligned} H_{G_I^0}(\mu,\alpha,L) &= H_{\Gamma_{\text{SAR}}} - \ln\Gamma(L-\alpha) \\ &+ (L-\alpha)\psi^{(0)}(L-\alpha) \\ &- (1-\alpha)\psi^{(0)}(-\alpha) + \ln(-1-\alpha) \\ &+ \ln\Gamma(-\alpha) - L, \end{aligned}$$

where $\psi^{(0)}(\cdot)$ is the digamma function.



$H_{G_I^0}$ converges to the $H_{\Gamma_{\text{SAR}}}$, $L=8$.

Non-parametric Entropy Estimation Approach

Vasicek introduced an non-parametric estimator in 1976:

$$\widehat{H}_{\mathsf{V}}(\boldsymbol{Z}) = \frac{1}{n}\sum_{i=1}^{n} \ln\left[\frac{n}{2m}\left(Z_{(i+m)} - Z_{(i-m)}\right)\right],$$

where $Z_{(i+m)} - Z_{(i-m)}$ is the m-spacing and $Z_{(1)} \leqslant Z_{(2)} \leqslant \ldots \leqslant Z_{(n)}$ are the **order statistics.**

Non-parametric Entropy Estimation Approach

Vasicek introduced an non-parametric estimator in 1976:

$$\widehat{H}_{\mathsf{V}}(\boldsymbol{Z}) = \frac{1}{n}\sum_{i=1}^{n} \ln\left[\frac{n}{2m}\left(Z_{(i+m)} - Z_{(i-m)}\right)\right],$$

where $Z_{(i+m)} - Z_{(i-m)}$ is the m-spacing and $Z_{(1)} \leqslant Z_{(2)} \leqslant \ldots \leqslant Z_{(n)}$ are the **order statistics.**

Non-parametric Entropy Estimation Approach

Vasicek introduced an non-parametric estimator in 1976:

$$\widehat{H}_{\mathsf{V}}(\boldsymbol{Z}) = \frac{1}{n}\sum_{i=1}^{n} \ln\left[\frac{n}{2m}\left(Z_{(i+m)} - Z_{(i-m)}\right)\right],$$

where $Z_{(i+m)} - Z_{(i-m)}$ is the m-spacing and $Z_{(1)} \leqslant Z_{(2)} \leqslant \ldots \leqslant Z_{(n)}$ are the **order statistics.**

We consider superior adaptations:

- [Ebrahimi et al., 1994]: $\widehat{H}_{\mathrm{E}}$.
- [Correa, 1995]: $\widehat{H}_{\mathrm{C}}$.
- [Al Omary, 2014]: $\widehat{H}_{\mathrm{AO}}$.

Non-parametric Entropy Estimation Approach

Vasicek introduced an non-parametric estimator in 1976:

$$\widehat{H}_{\mathsf{V}}(\boldsymbol{Z}) = \frac{1}{n}\sum_{i=1}^{n} \ln\left[\frac{n}{2m}\left(Z_{(i+m)} - Z_{(i-m)}\right)\right],$$

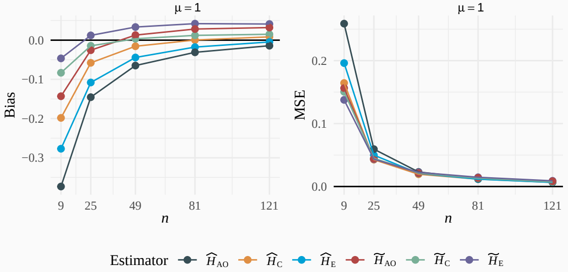
where $Z_{(i+m)} - Z_{(i-m)}$ is the m-spacing and $Z_{(1)} \leqslant Z_{(2)} \leqslant \ldots \leqslant Z_{(n)}$ are the **order statistics.**

We consider superior adaptations:

- [Ebrahimi et al., 1994]: $\widehat{H}_{\mathrm{E}}$.
- [Correa, 1995]: $\widehat{H}_{\mathrm{C}}$.
- [Al Omary, 2014]: $\widehat{H}_{\mathrm{AO}}$.

Enhanced Bootstrap Technique

$$\widetilde{H} = 2\widehat{\theta}(\boldsymbol{Z}) - \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_b(\boldsymbol{Z}^{(b)})$$



Bias and MSE of original and bootstrap versions, $L = 5$ and $\alpha = -10$.

Hypothesis Testing

Hypothesis Testing

We aim at testing the following hypotheses:

$$\begin{cases} \mathcal{H}_0 : \widetilde{H} = H_{\Gamma_{\mathsf{SAR}}} \\ \mathcal{H}_1 : \widetilde{H} \neq H_{\Gamma_{\mathsf{SAR}}}. \end{cases}$$

In other words, we verify the hypothesis that the data are fully-developed speckle with the following **test statistic**:

Hypothesis Testing

We aim at testing the following hypotheses:

$$\begin{cases} \mathcal{H}_0 : \widetilde{H} = H_{\Gamma_{\mathsf{SAR}}} \\ \mathcal{H}_1 : \widetilde{H} \neq H_{\Gamma_{\mathsf{SAR}}}. \end{cases}$$

In other words, we verify the hypothesis that the data are fully-developed speckle with the following **test statistic**:

$$S(\boldsymbol{Z}; L) = \widetilde{H} - \left[H_{\Gamma_{\mathsf{SAR}}}(L) + \ln \bar{\boldsymbol{Z}}\right].$$

The values should be around zero under the null hypothesis and far otherwise.

Hypothesis Testing

We aim at testing the following hypotheses:

$$\begin{cases} \mathcal{H}_0 : \widetilde{H} = H_{\Gamma_{\mathsf{SAR}}} \\ \mathcal{H}_1 : \widetilde{H} \neq H_{\Gamma_{\mathsf{SAR}}}. \end{cases}$$

In other words, we verify the hypothesis that the data are fully-developed speckle with the following **test statistic**:

$$S(\boldsymbol{Z}; L) = \widetilde{H} - \left[H_{\Gamma_{\mathsf{SAR}}}(L) + \ln \bar{\boldsymbol{Z}}\right].$$

The values should be around zero under the null hypothesis and far otherwise.

Hypothesis Testing

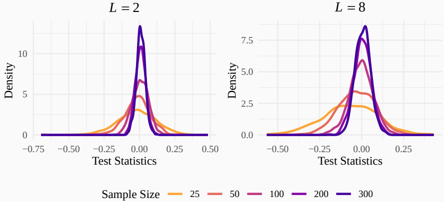
We aim at testing the following hypotheses:

$$\begin{cases} \mathcal{H}_0 : \widetilde{H} = H_{\Gamma_{\mathsf{SAR}}} \\ \mathcal{H}_1 : \widetilde{H} \neq H_{\Gamma_{\mathsf{SAR}}}. \end{cases}$$

In other words, we verify the hypothesis that the data are fully-developed speckle with the following **test statistic**:

$$S(\boldsymbol{Z}; L) = \widetilde{H} - \left[H_{\Gamma_{\mathsf{SAR}}}(L) + \ln \bar{\boldsymbol{Z}}\right].$$

The values should be around zero under the null hypothesis and far otherwise.



Empirical density under the null hypothesis.

Hypothesis Testing

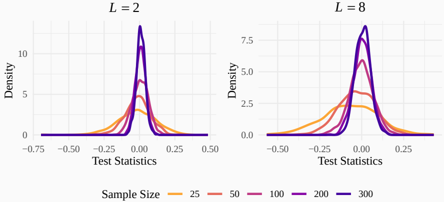
We aim at testing the following hypotheses:

$$\begin{cases} \mathcal{H}_0 : \widetilde{H} = H_{\Gamma_{\mathsf{SAR}}} \\ \mathcal{H}_1 : \widetilde{H} \neq H_{\Gamma_{\mathsf{SAR}}}. \end{cases}$$

In other words, we verify the hypothesis that the data are fully-developed speckle with the following **test statistic**:

$$S(\boldsymbol{Z}; L) = \widetilde{H} - \left[H_{\Gamma_{\mathsf{SAR}}}(L) + \ln \bar{\boldsymbol{Z}}\right].$$

The values should be around zero under the null hypothesis and far otherwise.



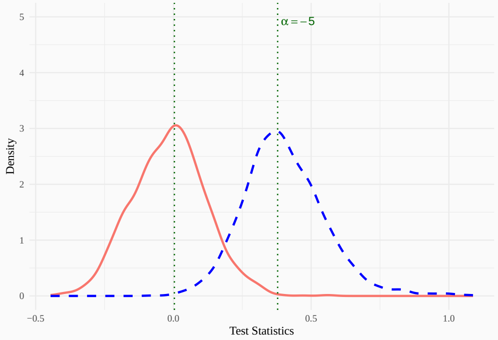
Empirical density under the null hypothesis.

L	n	Mean	SD	Var	SK	EK	p-value
2	25	−0.0199	0.1370	0.0188	−0.3492	1.2320	0.0586
	50	−0.0145	0.0896	0.0080	−0.1367	0.4826	0.0418
	100	0.0064	0.0562	0.0032	−0.0623	−0.1617	0.3938
	200	0.0067	0.0361	0.0013	0.0199	−0.0305	0.9273
	300	0.0093	0.0309	0.0010	0.0119	0.3522	0.1293
8	25	−0.0696	0.1623	0.0264	−0.1096	0.0158	0.1658
	50	−0.0301	0.1082	0.0117	−0.1741	0.0531	0.2961
	100	−0.0025	0.0711	0.0051	−0.0408	0.3446	0.2279
	200	0.0046	0.0520	0.0027	−0.0873	0.0466	0.5663
	300	0.0070	0.0448	0.0020	0.0818	0.0306	0.5431

Descriptive analysis of $S(\boldsymbol{Z}; L)$ to verify the normality of the data with $L \in \{2, 8\}$ and $\mu = 1$.

Hypothesis Testing

Empirical densities



Empirical distributions under the null (centered around zero) and alternative hypotheses (e.g., $\alpha = -5$, showing a shift from zero), with $\mu = 1$ and $L = 8$.

Size and Power of the proposed test.

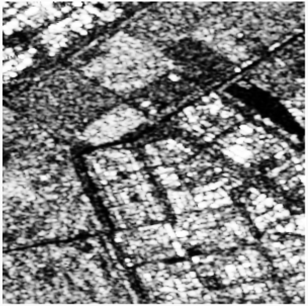
		Size			Power		
L	n	1%	5%	10%	1%	5%	10%
5	25	0.016	0.071	0.117	0.564	0.672	0.745
	50	0.014	0.052	0.113	0.640	0.754	0.806
	100	0.010	0.057	0.103	0.738	0.838	0.885
	200	0.010	0.054	0.092	0.807	0.937	0.963
	300	0.007	0.058	0.107	0.850	0.961	0.988
8	25	0.024	0.076	0.129	0.747	0.844	0.838
	50	0.015	0.055	0.104	0.857	0.919	0.939
	100	0.006	0.052	0.099	0.958	0.984	0.985
	200	0.014	0.043	0.109	0.989	0.999	1.000
	300	0.014	0.041	0.106	1.000	1.000	1.000

Application to Actual Data

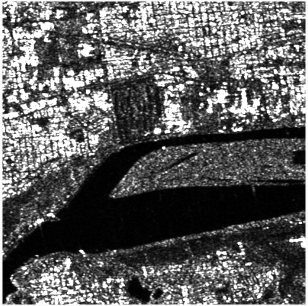
Application to Actual Data

Images acquired by the Sentinel-1A satellite operating in the C band, with VV polarization, intensity format, and $L = 5$ nominal looks.

Both images contain homogeneous and heterogeneous zones.



(a) Flevoland

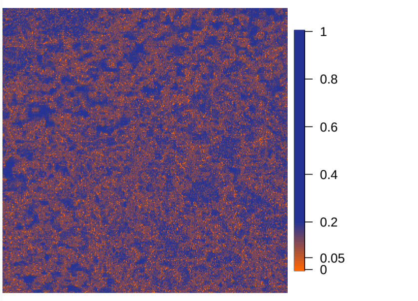


(b) Ottawa

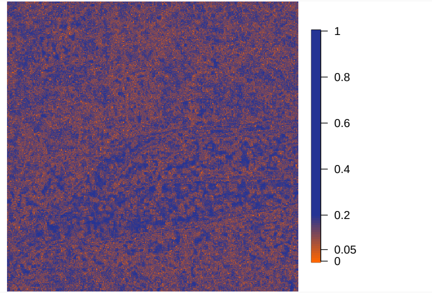
SAR images, 300×300 and 512×512 pixels, respectively.

Application to Actual Data

The color table maps all above 0.05 into dark blue (no evidence to reject the fully-developed speckly hypothesis), and those below 0.05 into a gradient between blue and orange (evidence to reject the hypothesis).



(a) Flevoland



(b) Ottawa

Heatmap for a threshold of 0.05 of the p-values, generated after applying the proposed using 7×7 local sliding windows.

Conclusions & Future Perspectives

Conclusions & Future Perspectives

- We proposed a testing procedure to distinguish between fully-developed speckle and heterogeneous clutter.
- This tool is equipped with a test statistic based on a bootstrap-improved estimator of the Shannon entropy.
- Applying the test to actual SAR data yielded promising results, effectively distinguishing different regions in the image and demonstrating the test's potential to differentiate between homogeneous and heterogeneous zones.

Currently, we are working on detecting texturelessness in SAR images. To this end, we propose novel hypothesis tests based on classical and variants of the coefficient of variation.

Q&A

Thank you for your attention

Contact
Alejandro C. Frery
alejandro.frery@vuw.ac.nz

Scan to connect