

## Modelo de Regressão Aplicado à Dados de Radar de Abertura Sintética

Willams B. F. da Silva,<sup>1</sup> Abraão D. C. Nascimento,<sup>2</sup> Francisco José A. Cysneiros.<sup>3</sup>  
Departamento de Estatística - Universidade Federal de Pernambuco, Recife, PE.

O Radar de Abertura Sintética (SAR) é uma ferramenta essencial para o sensoriamento remoto. Uma abordagem interessante para imagens SAR é o modelo multiplicativo, que considera o produto entre retroespalhamento do terreno e o efeito do ruído speckle na imagem SAR. É usual que esse modelo multiplicativo tenha como suposição a distribuição gama para o ruído *speckle*, porém, a metodologia de análise de variância para imagens proposta por Ludwig & Kurz (1997) [1] aborda o tema supondo normalidade dos erros. Neste trabalho, relacionamos a família de distribuição log-simétrica proposta por Vanegas & Paula (2015) [2] com o modelo multiplicativo de imagens SAR e abordamos a modelagem como um problema de análise de variância. Essa mudança distribucional é crucial para os atributos dessas imagens, já que a distribuição gaussiana não é adequada, para intensidade e amplitude ambas positivas.

A abordagem multiplicativa define cada pixel como o produto de duas variáveis aleatórias independentes: o retroespalhamento do terreno e o ruído de speckle. Este estudo foca em imagens de intensidade SAR. Considere uma variável aleatória  $Z$  contínua e simétrica, com parâmetros  $\mu$  e  $\phi$ , então  $T = \exp(Z)$  tem função de densidade de probabilidade (f.d.p.) dada por

$$f_T(t) = \frac{g \left[ \log^2(t/\eta)^{\frac{1}{\phi}} \right]}{t\sqrt{\phi}} \mathbb{I}_{\mathbb{R}_+}(t), \quad (1)$$

em que  $\eta = \exp(\mu)$ ,  $\mathbb{I}_{\mathbb{R}_+}(\cdot)$  é a função indicadora no conjunto  $\mathbb{R}_+$  e  $g(\cdot)$  é uma função positiva em  $(0, \infty)$  tal que  $\int_0^\infty x^{-1/2} g(x) dx = 1$ . Sendo  $T \sim LS(\eta, \phi, g(\cdot))$ . O modelo de regressão para esta família pode ser escrito como

$$T_k = \eta_k \epsilon_k^{\sqrt{\phi_k}}, \quad (2)$$

para  $k = 1, 2, \dots, n$ , onde  $\eta_k > 0$  e  $\phi > 0$  representam a mediana e a assimetria (dispersão relativa) da distribuição de  $T_k$ , e  $\epsilon_1, \epsilon_2, \dots, \epsilon_n$  são um conjunto de variáveis aleatórias log-simétrica padrão, ou seja,  $\epsilon_k \sim LS(1, 1, g(\cdot))$ . Para este trabalho, usamos o modelo de dois fatores que correspondem aos efeitos das linhas e colunas da imagem. O janelamento usado é de  $5 \times 5$ , isto é, dividimos a imagem SAR em blocos de 25 observações por bloco, totalizando para nossa aplicação 6161 blocos na imagem. Para o nosso modelo, as equações da função de ligação ficam

$$\begin{cases} \log(\eta_k) &= \mu + \alpha_i + \beta_j, \\ \log(\phi_k) &= \gamma, \end{cases} \quad (3)$$

para  $i = 1, \dots, l$  e  $j = 1, \dots, m$ , onde  $l$  e  $m$  são os tamanhos das linhas e colunas do janelamento, respectivamente, sendo  $\sum_{i=1}^l \alpha_i = 0$  e  $\sum_{i=1}^m \beta_j = 0$ . A estrutura de dois fatores refere apenas à mediana da distribuição; o parâmetro de assimetria não possui estrutura.

<sup>1</sup>willams.bfsilva@ufpe.br

<sup>2</sup>abraao@de.ufpe.br

<sup>3</sup>cysneiros@de.ufpe.br

Para ilustrar nossa metodologia, apresentamos uma imagem SAR obtida pelo satélite Sentinel-1 em abril de 2023, com fator *multilook* de 3 no eixo de alcance e 1 no azimutal. A região é próxima à cidade de Japaratinga, incluindo o Rio Manguaba em Alagoas, Brasil. A Figura 1(a) exibe os valores de intensidade do canal VH de polarização e as Figuras 1(b)–(d) mostram as estimativas de  $\eta$  que é frequentemente chamado de imagens SAR filtradas ou preditas a partir de diferentes distribuições. Fig 1(b) para a distribuição log-t-Student com 5 graus de liberdade, Fig 1(c) para a distribuição log-normal e a Fig 1(d) para a distribuição normal. Observa-se que as estimativas supondo a distribuição gaussiana levam à valores negativos para a intensidade, indicando um valor fora da natureza da variável. Adicionalmente, valores maiores que 0,312 estão presentes na imagem real, mas não na estimada pelo modelo gaussiano. Os valores para os erros quadrático médio (RMSE) foram de 0,0178 para o modelo normal, de 0,0147 para a distribuição log-normal e de 0,0175 para log-t-Student com 5 graus de liberdade, indicando um melhor ajuste para o modelo log-normal que os demais.

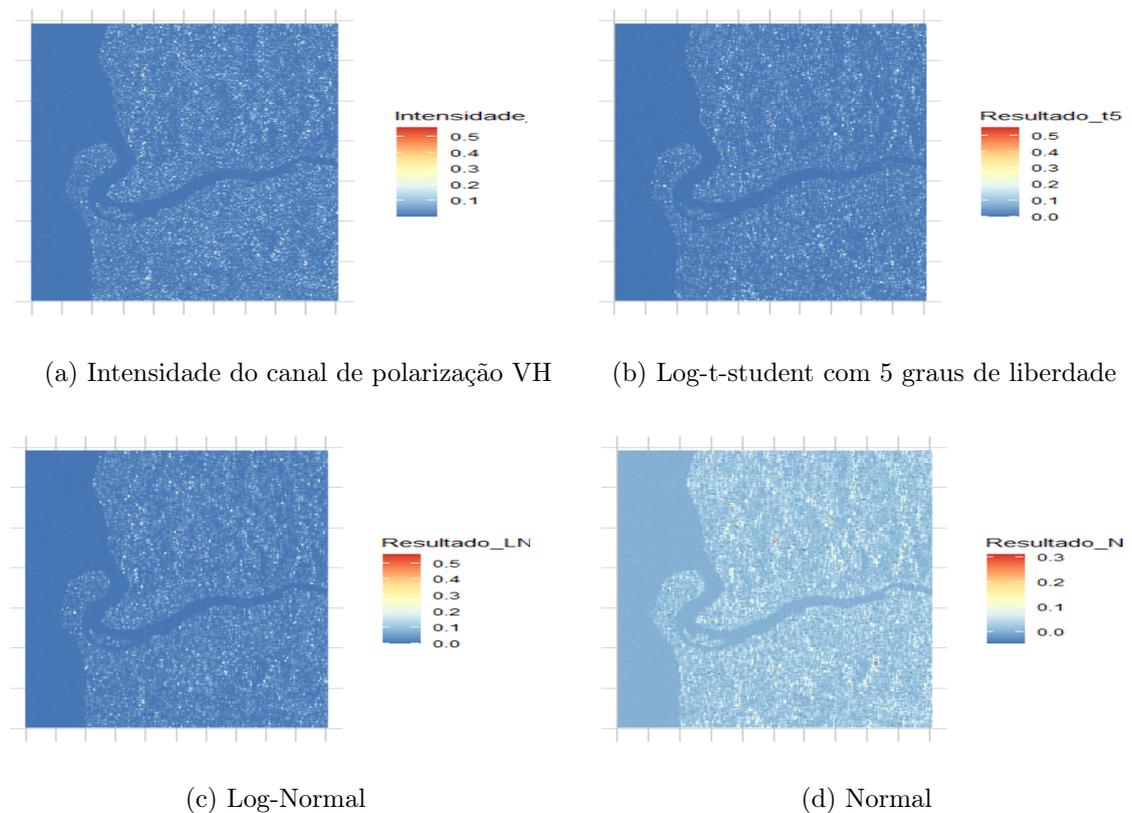


Figura 1: (a) Imagem real SAR e (b–d) imagens preditas sob diferentes distribuições. Fonte: Os autores.

## Referências

- [1] L. Kurz e M. H. Benteftifa. **Analysis of Variance in Statistical Image Processing**. Cambridge, United Kingdom: Cambridge University Press, 1997. DOI: 10.1017/cbo9780511530166.
- [2] L. H. Vanegas e G. A. Paula. “A semiparametric approach for joint modeling of median and skewness”. Em: **TEST** 24.1 (2015), pp. 110–135. ISSN: 1133-0686. DOI: 10.1007/s11749-014-0401-7.