



Universidade Federal do Rio Grande do Sul  
Instituto de Matemática e Estatística  
Programa de Pós-Graduação em Estatística

# **Modelo Rayleigh Escore Autorregressivo Generalizado para Interpretação de Dados de Radar de Abertura Sintética**

Miguel Roberto Peña Ramírez

Porto Alegre, Abril de 2023.



### CIP - Catalogação na Publicação

Peña-Ramírez, Miguel Roberto  
Modelo Rayleigh Escore Autorregressivo Generalizado  
para Interpretação de Dados de Radar de Abertura  
Sintética / Miguel Roberto Peña-Ramírez. -- 2023.  
49 f.  
Orientadora: Renata Rojas Guerra.

Coorientador: Fábio Mariano Bayer.

Dissertação (Mestrado) -- Universidade Federal do  
Rio Grande do Sul, Instituto de Matemática e  
Estatística, Programa de Pós-Graduação em Estatística,  
Porto Alegre, BR-RS, 2023.

1. Distribuição Rayleigh. 2. Imagens SAR. 3.  
Modelos GAS. 4. Parametrização na média. 5.  
Sensoriamento remoto. I. Rojas Guerra, Renata, orient.  
II. Bayer, Fábio Mariano, coorient. III. Título.



Dissertação submetida por Miguel Roberto Peña Ramírez como requisito parcial para a obtenção do título de Mestre em Estatística pelo Programa de Pós-Graduação em Estatística da Universidade Federal do Rio Grande do Sul.

**Orientador(a):**

Profa. Dra. Renata Rojas Guerra

**Co-orientador(a):**

Prof. Dr. Fábio Mariano Bayer

**Comissão Examinadora:**

Prof. Dr. Abraão David Costa do Nascimento (UFPE)

Profa. Dra. Bruna Gregory Palm (Blekinge Institute of Technology)

Profa. Dra. Taiane Schaedler Prass (PPGEst - UFRGS)

Data de Apresentação: 3 de abril de 2023



*À minha família,  
especialmente aos meus pais,  
a quem devo tudo o que sou.*





# AGRADECIMENTOS

*Agradeço primeiro à Profa. Dra. Renata Rojas, pela orientação, colaboração constante, dedicação e paciência, e por conduzir este trabalho de pesquisa no caminho certo. Também ao Prof. Dr. Fábio Bayer, pela colaboração na coorientação, por suas sugestões constantes, por me dar a orientação inicial para a escolha do tema de pesquisa e por me fornecer as bases iniciais para o desenvolvimento deste trabalho.*

*Igualmente agradeço ao Programa de Pós-Graduação em Estatística (PPGEst) por me dar a oportunidade de ser aluno de tão importante programa e universidade como a Universidade Federal do Rio Grande do Sul (UFRGS), da qual sinto orgulho de ter sido aluno. Assim como a todos os professores do PPGEst, pelo acompanhamento durante o mestrado e pelo conhecimento compartilhado, bem como pela excelência acadêmica prestada.*

*Aos professores membros da banca pela disponibilidade e pelas considerações e valiosas contribuições. Aos colaboradores do PPGEst pelo apoio e assistência oportuna durante minha permanência no programa.*

*À minha família pelo constante apoio, especialmente ao meu irmão, Fernando, que foi quem me motivou e incentivou para realizar o mestrado.*

*A todas as outras pessoas que durante este tempo contribuíram para tornar este trabalho uma realidade.*



## RESUMO

Retornos em amplitude de imagens de radar de abertura sintética (SAR, do inglês *synthetic aperture radar*) apresentam comportamento assimétrico e valores estritamente positivos, sendo adequadamente caracterizados pela distribuição Rayleigh em regiões imageadas homogêneas. Além disso, imagens SAR comumente apresentam dependência espacial entre os pixes adjacentes devido à maneira como os dados SAR são adquiridos e processados. É neste contexto que se insere o presente trabalho, no qual é proposto o modelo Rayleigh escore autorregressivo generalizado (Ray-GAS, do inglês *Rayleigh generalized autoregressive score*). O Ray-GAS é um modelo dinâmico útil para a interpretação de dados SAR derivado da estrutura da classe de modelos escore autorregressivo generalizados (GAS, do inglês *generalized autoregressive score*). Para o desenvolvimento do Ray-GAS, assume-se que a média condicional da distribuição Rayleigh é um parâmetro variante no tempo (ao longo dos índices da imagem). Na presente dissertação, uma revisão bibliográfica é feita e, logo após, o modelo proposto é apresentado no artigo intitulado "*The Rayleigh Generalized Autoregressive Score Model for SAR Data Interpretation*", sendo o principal produto deste trabalho. Dentre os desenvolvimentos realizados, tem-se o estudo de aspectos de estimação pontual por meio do método de máxima verossimilhança condicional e de ferramentas para análise de diagnóstico e predição, com base no novo modelo da classe GAS. Além disso, são realizadas simulações de Monte Carlo para avaliar o desempenho dos estimadores apresentados. Por fim, a metodologia proposta é aplicada em dados medidos provenientes de imagens SAR do sistema CARABAS II em regiões de floresta e lago. Os resultados evidenciam a potencialidade do modelo Ray-GAS na interpretação do comportamento estocástico e a obtenção de predições de retornos em amplitude de imagens SAR.



## ABSTRACT

Amplitude returns from synthetic aperture radar (SAR) images show asymmetric behavior and strictly positive values, adequately characterized by the Rayleigh distribution in homogeneous imaged regions. In addition, SAR images commonly show spatial dependence between adjacent pixels due to how SAR data is acquired and processed. The present work is inserted in this context, in which the Rayleigh generalized autoregressive score (Ray-GAS) model is proposed. Ray-GAS is a usable dynamic model for interpreting SAR data derived from the structure of the generalized autoregressive score (GAS) class of models. For the development of Ray-GAS, the conditional mean of the Rayleigh distribution is assumed to be time-varying (on the image indices). In this dissertation, a literature review is carried out, followed by the presentation of the model proposed in the article entitled "*The Rayleigh Generalized Autoregressive Score Model for SAR Data Interpretation*," which is the main product of this work. Among the developments, the point estimation of the parameters is carried out utilizing the conditional maximum likelihood method and tools for diagnostic and prediction analysis based on the new GAS class model. In addition, Monte Carlo simulations are performed to evaluate the performance of the presented estimators. Finally, the proposed methodology is applied to measured data from SAR images of the CARABAS II system in forest and lake regions. The results show the potential of the Ray-GAS model in interpreting the stochastic behavior and obtaining predictions of returns in the amplitude of SAR images.



---

# ÍNDICE

---

<b>1</b>	<b>Introdução</b>	<b>2</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>6</b>
2.1	Imagens SAR . . . . .	6
2.2	Distribuição Rayleigh Reparametrizada . . . . .	7
2.3	A Especificação Básica dos Modelos GAS . . . . .	10
<b>3</b>	<b>Desenvolvimento</b>	<b>14</b>
<b>4</b>	<b>Conclusões e trabalhos futuros</b>	<b>28</b>
	<b>Apêndices</b>	<b>32</b>
<b>A</b>	<b>Descrição imagem CARABAS-II</b>	<b>33</b>

---

# CAPÍTULO 1

## INTRODUÇÃO

---

O sensoriamento remoto envolve a captura de radiações eletromagnéticas da superfície terrestre, sendo que os sistemas de sensoriamento remoto podem ser classificados em sistemas passivos ou ativos (Moreira, 2005). Os sistemas passivos requerem uma fonte de energia externa que emita radiação, por exemplo, sensores óticos requerem luz solar. Por outro lado, sistemas ativos são aqueles que emitem sua própria fonte de radiação eletromagnética e usam transmissores que emitem ondas para a superfície da terra. Assim, parte destas ondas são refletidas novamente e recebidas pelo sensor para registrar as informações capturadas. Dentre os sensores ativos mais utilizados destaca-se o sistema de radar de abertura sintética (SAR, do inglês *synthetic aperture radar*). Sua principal vantagem em relação a outros sistemas, tais como os óticos, é que permite a obtenção de dados mesmo sob condições adversas.

De acordo com Poulain et al. (2011), os sensores SAR independem das condições climáticas e podem fornecer, com eficácia, informações a qualquer hora do dia ou da noite, ou em áreas nubladas, enquanto os sensores óticos dependem destes fatores. Segundo Meyer et al. (2016), as imagens de sensoriamento remoto são indispensáveis para o monitoramento contínuo da cintura equatorial, onde a cobertura constante de nuvens e chuvas fortes dificulta o funcionamento dos sistemas de satélites que operam em frequências óticas ou térmicas. Mas os sistemas SAR, devido a sua maior coerência de tempo e penetração mais profunda, proporcionam melhor desempenho. Eineder et al. (2010) demonstraram que em alguns sensores SAR as imagens apresentam precisões de geolocalização absolutas dentro de poucos decímetros, sendo outra vantagem frente aos sensores óticos de alta resolução que requerem pontos de controle na terra para alcançar precisões similares. Assim, a imagem ótica revela o que os olhos humanos podem observar na realidade, enquanto a imagem SAR ajuda a visualizar o mundo eletromagnético (Merkle et al., 2018).

Na geração de imagens SAR, a amplitude é uma medida que representa a magnitude da resposta complexa do sinal do radar. A amplitude do sinal recebido por um sensor SAR depende da geometria do sistema de radar, da polarização do sinal, da distância entre o sensor e o alvo, das características do alvo e das propriedades do ambiente entre o sensor e o alvo. Também depende do número de visadas da imagem, seja *single-look* ou *multi-look*.

Apesar das inúmeras vantagens, o processamento e a interpretação dos sistemas SAR estão sujeitos aos efeitos aleatórios de uma interferência granular denominada ruído *speckle* (Oliver and Quegan, 2004). A presença de ruído *speckle* pode degradar as imagens SAR, levando à perda de informações



(Kuruoglu and Zerubia, 2004). Neste contexto, os modelos estatísticos tornam-se essenciais para a descrição e interpretação de imagens SAR (Gao, 2010) e podem ser utilizados para diferentes finalidades. Por exemplo, Liu et al. (2019) baseiam-se em abordagens estatísticas para resolver problemas de classificação. Bayer et al. (2020a), Palm et al. (2022a) e Sagrillo et al. (2023) consideram tais métodos no contexto de detecção de anomalias. Novas distribuições de probabilidade tem sido propostas como descritores de imagens SAR, como pode ser visto em Sagrillo et al. (2022), Cintra et al. (2012), Cordeiro et al. (2019) e Vasconcelos et al. (2020). Além disso, é também possível encontrar aplicações com objetivos de previsão de pixels (Almeida-Junior and Nascimento, 2021), segmentação (Meng et al., 2019) e detecção de bordas (Chen et al., 2019).

O uso de modelos estatísticos clássicos, foram derivados com base no modelo speckle. Imagens observadas em intensidade single-look podem ser modeladas usando a distribuição exponencial negativa, em amplitude single-look podem ser modeladas usando uma distribuição Rayleigh, enquanto a distribuição Gamma é frequentemente usada para modelar a amplitude de pixels em imagens multi-look (Gao, 2010). A distribuição Rayleigh é comumente usada para modelar a amplitude das imagens SAR em áreas homogêneas porque essas áreas tendem a ter uma retroespalhamento mais uniforme e regular, que se encaixa bem nessa distribuição de probabilidade (Oliver and Quegan, 2004). Generalizações e aplicações da distribuição Rayleigh em imagens SAR podem ser verificadas em Kuruoglu and Zerubia (2004), Li et al. (2010), Peng and Zhao (2013) e Yue et al. (2020). Contudo, no contexto de uma imagem SAR, presumir que o parâmetro de escala na distribuição Rayleigh é fixo significa não haver variação espacial nas amplitudes dos pixels, o que pode ser restritivo em diversos problemas práticos no âmbito da modelagem de imagens SAR.

Torna-se importante ressaltar que a dependência espacial entre pixels é uma característica inerente à natureza de como os dados são adquiridos e processados. Além de utilizar uma distribuição de probabilidade adequada, é importante considerar modelos com estruturas que considerem esta dependência. Assim, uma linha de pesquisa que tem sido explorada é a proposição de modelos que generalizam as suposições de independência por meio de diferentes estruturas. Algumas contribuições mais recentes, sugerem como base uma reparametrização em termos da média, a qual é introduzida em Palm et al. (2019). A vantagem desta abordagem é sua interpretação como medida de tendência central, sendo tipicamente útil não apenas para obter estruturas dinâmicas, mas também no contexto de análise de regressão.

Em Palm et al. (2019) é proposto um modelo de regressão no qual a média dos sinais Rayleigh-distribuídos é influenciada por regressores. No contexto de estruturas dinâmicas, pode-se referenciar o modelo Rayleigh autorregressivo de médias móveis (RARMA, do inglês *Rayleigh autoregressive moving average*) (Bayer et al., 2020b). Além disso, uma generalização do RARMA é também proposta para acomodar estruturas de dependência bidimensionais (Palm et al., 2022a). No entanto, é fundamental destacar que outras estruturas dinâmicas podem ser desenvolvidas para melhorar a interpretação de dados de imagem SAR. Assim, a classe de modelos escore autorregressivo generalizado (GAS, do inglês *generalized autoregressive score*) surge como uma alternativa.

Originalmente proposta por Creal et al. (2008) e Creal et al. (2013), a classe GAS abrange modelos dinâmicos com parâmetros variantes no tempo cujo mecanismo de atualização é uma função do vetor escore ponderado e de seus valores passados. Dentre as vantagens no uso desta modelagem, Creal et al. (2008) destaca sua capacidade de explorar a estrutura completa da densidade no mecanismo de atualização dos parâmetros, e não apenas a média ou momentos de ordem superior.

Desde sua proposta, os modelos GAS têm sido de grande utilidade para o ajuste de séries temporais econômicas e financeiras. O foco nestas áreas de aplicação pode justificar-se pelo fato de que alguns modelos clássicos destas áreas são casos particulares da classe GAS. Alguns exemplos são os modelos autorregressivos condicionalmente heterocedásticos (Engle and Bollerslev, 1986), modelos de erro multiplicativo (Engle, 2002), e os modelos autorregressivos de intensidade e duração condicional (Engle and Russell, 1998). Além desses modelos clássicos, pode-se referenciar Alanya-Beltran (2022), Fuentes et al. (2023) e Folqué et al. (2021) como algumas recentes contribuições a esta discussão. Por outro lado, em Lucas and Koopman (2023), pode-se ver que as vantagens da classe GAS não foram exploradas no contexto do sensoriamento remoto ou, mais especificamente, para modelagem e predição de imagens SAR.

### *Objetivo*

Este trabalho tem por objetivo propor um modelo para interpretação de dados SAR onde a estrutura do parâmetro da distribuição condicional Rayleigh é modelada com uma estrutura GAS. Além da proposição do modelo, são introduzidas a teoria de estimação puntual dos parâmetros do modelo, ferramentas para análise de diagnóstico e uma expressão para obtenção de predições um passo à frente. A teoria proposta é avaliada numericamente via simulações de Monte Carlo. O modelo introduzido é comparado em termos de desempenho de predição com o modelo RARMA, obtendo, em geral, melhores resultados.

### *Novidades do trabalho*

Este trabalho pretende contribuir no sensoriamento remoto através da proposta de um modelo da classe GAS com base na distribuição Rayleigh reparametrizada na média. Uma vez que se considera um modelo clássico para retornos em amplitude de imagens como distribuição preditiva, espera-se que a proposta seja uma boa alternativa para a modelagem e predição de imagens SAR, mantendo as vantagens da estrutura dinâmica da classe GAS. Desta forma, pretende-se fornecer evidências de que o modelo proposto pode contribuir para a compreensão do comportamento estocástico e prever valores de retornos em amplitude de imagens SAR, sendo uma alternativa eficaz ao modelo RARMA.

### *Suporte computacional*

Todos os aspectos metodológicos desenvolvidos nesta dissertação, bem como os estudos numéricos simulados e aplicados, foram implementados em linguagem de programação R na versão 4.2.2.2 (R Core Team, 2022). Os códigos usados estão disponíveis em <https://github.com/miguerpe/Ray-GAS>.

### *Organização do trabalho*

A presente dissertação está organizada da seguinte forma. No Capítulo 2, é apresentada uma revisão bibliográfica abrangendo alguns temas relevantes para o desenvolvimento do problema de interesse. Inclui uma introdução à geração de imagens SAR, uma descrição das propriedades da distribuição Rayleigh considerando a reparametrização em termos da média, sendo descritos os modelos de regressão Rayleigh e o modelo RARMA. Além disso, a especificação geral dos modelos GAS é apresentada neste mesmo capítulo. No Capítulo 3, é apresentado o artigo intitulado "*The Rayleigh Generalized Autoregressive Score Model for SAR Data Interpretation*", o qual introduz o modelo Ray-GAS, apresenta os aspectos metodológicos desenvolvidos e as simulações de Monte Carlo para evidenciar o potencial da proposta para a modelagem e predição de retornos em amplitude de imagens SAR mensuradas e simuladas. Finalmente, o Capítulo 4 apresenta as considerações finais desta dissertação, sendo destacadas algumas propostas para pesquisas futuras relacionadas a este trabalho.

---

## CAPÍTULO 2

# REVISÃO BIBLIOGRÁFICA

---

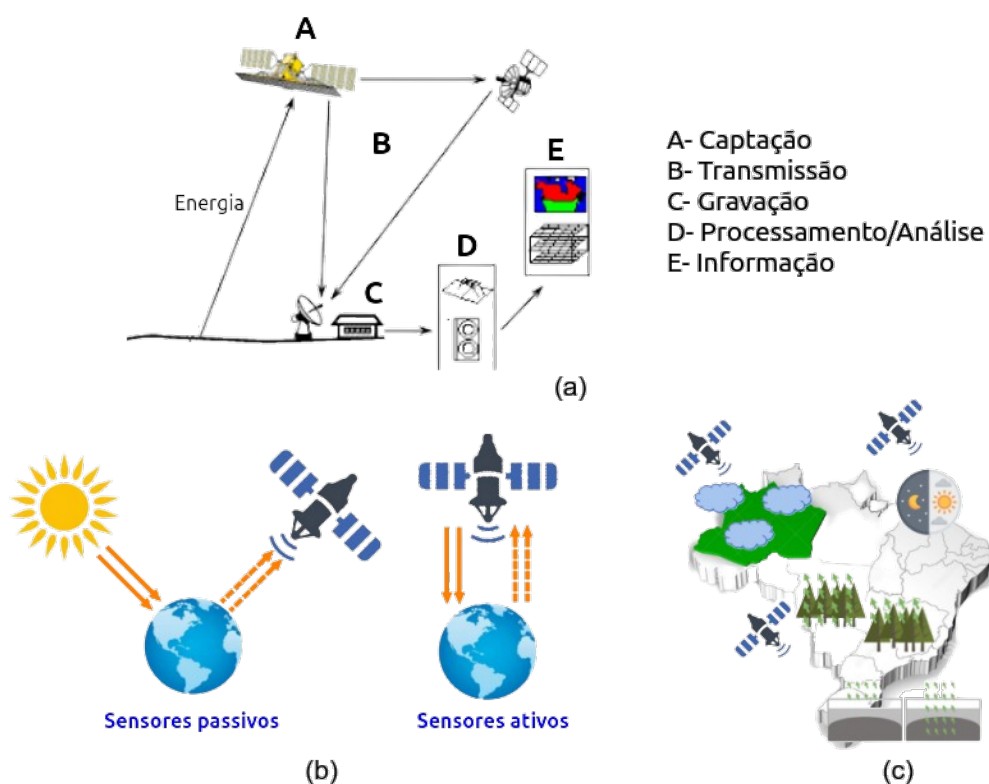
O presente capítulo é destinado à apresentação dos principais conceitos relacionados à construção metodológica desta dissertação. Inicialmente, uma descrição geral sobre imagens SAR é feita. Depois são fornecidas informações acerca da distribuição Rayleigh reparametrizada em termos da média e sua relevância para o desenvolvimento de modelos de regressão e séries temporais para a modelagem de imagens SAR. Em seguida, introduz-se a especificação geral dos modelos GAS, suas propriedades e algumas densidades preditivas que já foram desenvolvidas sob a estrutura desta classe.

### 2.1 *Imagens SAR*

Uma imagem SAR é construída a partir de dados obtidos por um radar de abertura sintética. O radar pode ser situado em um satélite ou em uma aeronave tripulada ou não tripulada. O radar emite pulsos de ondas eletromagnéticas que são refletidas na terra e nos objetos que encontram nela. Finalmente, esses sinais são captados por um sensor no mesmo radar. Assim, é possível obter informações detalhadas sobre a superfície terrestre e os objetos que nela possam existir. Uma das principais características dos sistemas SAR é sua capacidade de penetrar em diferentes superfícies, possibilitando detectar objetos que não são visíveis em um sistema óptico ou detectar pequenas alterações na superfície, tornando-se útil em muitas aplicações. A Figura 2.1 ilustra graficamente alguns desses conceitos. Uma importante referência sobre os princípios da imagens SAR é [Oliver and Quegan \(2004\)](#).

As imagens SAR são dadas pelas propriedades da cena da imagem. Quando o sinal é enviado, fatores como interação com a superfície do alvo, irregularidades e composição do alvo, entre outros fatores, afetam o sinal retornado. Assim, o sistema pode modificar, distorcer ou produzir incertezas nessas propriedades. Essas propriedades são determinadas pelo desenho do sistema ([Oliver and Quegan, 2004](#)). A Figura 2.2(a) mostra algumas das principais propriedades que determinam tais imagens. Esses fatores influenciam os mecanismos de dispersão e absorção das ondas. Assim, uma imagem SAR contém as características de dispersão da superfície de uma determinada área.

A imagem SAR é construída a partir da combinação de múltiplos pulsos de radar sucessivos emitidos em diferentes momentos e de diferentes locais. Medindo o eco de cada pulso que rebate na superfície e permitindo a obtenção de uma imagem da superfície da Terra. Cada pulso do radar refletido na cena imageada é registrado como um número complexo, contendo informações de amplitude e fase.



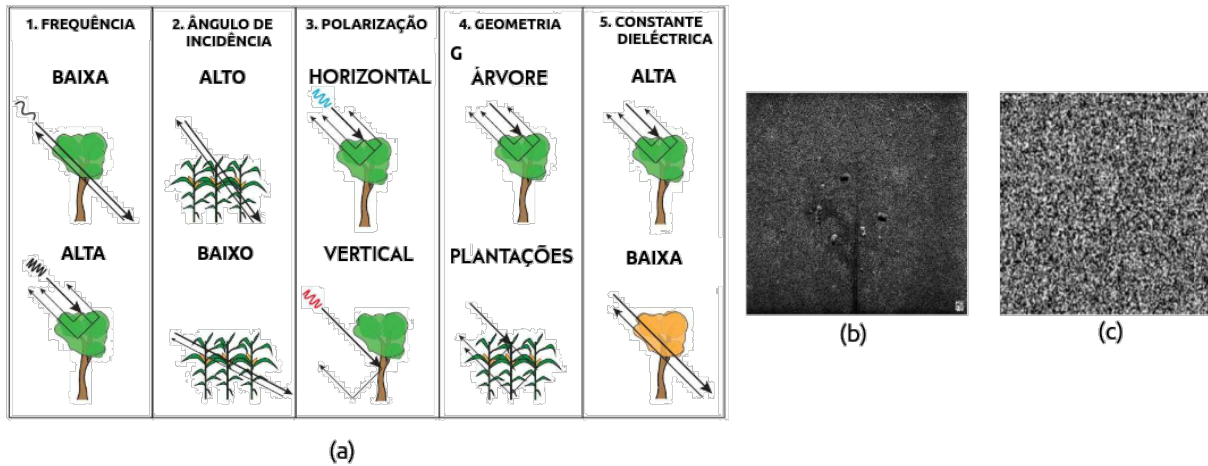
**Figura 2.1:** (a) Sensoriamento remoto. (b) Tipos de sensores. (c) Vantagens dos sistemas SAR.

Usualmente, para análises de imagens SAR, é considerada apenas a amplitude do sinal. A imagem de amplitude capturada estará em formato digital, representada por uma matriz de linhas e colunas, onde cada elemento é chamado de pixel. Assim, as informações coletadas por cada linha de pixels em uma imagem SAR podem ser consideradas como um sinal unidimensional, ou série temporal, que mostra a evolução da superfície da Terra ao longo da estrutura espacial. Ainda, quando a superfície é imageada, a interferência das ondas refletidas geram um ruído denominado speckle, o qual dificulta a interpretação das imagens SAR. O speckle provoca uma variação nos pixels causando um padrão granular na imagem (Lee and Pottier, 2017), conforme é mostrado na Figura 2.2.

Uma imagem SAR também pode ser gerada conforme o número de visadas. O procedimento consiste em imagear um mesmo pixel  $L$  vezes e gerar um único pixel por meio da média aritmética resultante (Lee and Pottier, 2017), sendo  $L$  o número de looks. Esse procedimento busca diminuir o efeito do ruído speckle. As imagens com  $L > 1$  são conhecidas como imagens de múltiplas visadas (multi-look) e no caso de  $L = 1$  a imagem é denominada de uma visada (single-look).

## 2.2 Distribuição Rayleigh Reparametrizada

A distribuição Rayleigh é comumente usada para modelar a amplitude em imagens SAR single-look (Gao, 2010). Quando um sinal é enviado a um alvo, a energia refletida é espalhada em várias direções e sofre influência aleatória de diversos fatores. Assim, quando o sinal é recebido pelo sistema SAR é usual considerá-lo como uma variável aleatória. A variável aleatória complexa pode ser decomposta



**Figura 2.2:** (a) Fatores imagens. SAR (b) ruído *Speckle* em uma imagem SAR. (c) Padrão do *speckle*.

em seus componentes ortogonais (real e imaginário) no plano complexo. Se presumirmos que esses componentes são independentes e seguem distribuição Gaussiana com média zero (devido à aleatoriedade da retrodifusão) e possuem a mesma variância  $\sigma^2$ , a amplitude desse sinal, ou seja, a raiz quadrada da soma dos quadrados desses componentes, segue distribuição Rayleigh (Soumekh, 1999). Esta interpretação faz com que a distribuição Rayleigh seja muito popular no contexto de modelagem e processamento de imagens. Assim, sob essas premissas, os valores de amplitude de dados SAR seguem a distribuição Rayleigh, com densidade dada por

$$f_Y(y; \sigma) = \frac{y}{\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right), \quad (2.1)$$

em que  $y > 0$ , e  $\sigma > 0$  é um parâmetro de escala. Sendo  $Y$  uma variável aleatória com função de densidade de probabilidade dada em (2.1), a média e a variância deste modelo são dadas, respectivamente, por

$$E(Y) = \sigma \sqrt{\frac{\pi}{2}} \quad \text{e} \quad \text{Var}(Y) = \sigma^2 \left(\frac{4 - \pi}{2}\right).$$

Além de ser utilizada em diversas aplicações da área de sensoriamento remoto, modificações e/ou generalizações da distribuição Rayleigh têm sido estudadas com o intuito de adicionar flexibilidade, ou uma melhor interpretabilidade ao modelo original. Neste contexto, um importante avanço é a reparametrização em termos da média, que foi introduzida por Palm et al. (2019) para formular o modelo de regressão Rayleigh. Sob esta proposta, toma-se  $\sigma = \mu/(\pi/2)^{(1/2)}$  e substitui-se essa expressão na Equação (2.1). Assim, uma variável aleatória  $Y$  tem a distribuição Rayleigh com parametrização na média se sua função de densidade de probabilidade é dada por

$$f_Y(y; \mu) = \frac{\pi y}{2\mu^2} \exp\left(-\frac{\pi y^2}{4\mu^2}\right), \quad (2.2)$$

em que  $y > 0$  e  $\mu > 0$  é a média de  $Y$ .

Considerando a proposta de Palm et al. (2019), as funções de distribuição acumulada e quantílica de  $Y$  são dadas por

$$F_Y(y; \mu) = 1 - \exp\left(-\frac{\pi y^2}{4\mu^2}\right) \quad \text{e} \quad Q_Y(u; \mu) = 2\mu \sqrt{\frac{-\log(1 - u)}{\pi}},$$

respectivamente. Claramente, neste caso, se tem também que a média de  $Y$  reduz-se a

$$\mathbb{E}(Y) = \mu$$

e sua variância é expressa por

$$\text{Var}(Y) = \mu^2 \left( \frac{4}{\pi} - 1 \right).$$

A função de quantílica é particularmente útil para a geração de valores pseudo-aleatórios da distribuição Rayleigh pelo método da inversão. Isso nos permitirá realizar simulações de Monte Carlo sob diversos cenários.

[Palm et al. \(2019\)](#) também propõem um modelo de regressão com base nesta reparametrização. Assim, sejam  $Y[1], Y[2], \dots, Y[n]$  variáveis aleatórias independentes, em que cada variável segue a densidade Rayleigh dada pela Equação (2.2) com média  $\mu[t]$ ,  $\mu[t]$ ,  $t = 1, 2, \dots, n$ . O modelo de regressão Rayleigh é obtido considerando um preditor linear  $\eta[t]$  para a média de  $Y[t]$  fornecido pela seguinte estrutura linear

$$\eta[t] = g(\mu[t]) = \sum_{i=1}^r \beta_i x_i[t], \quad t = 1, 2, \dots, n, \quad (2.3)$$

em que  $r < n$  é o número de covariáveis do modelo,  $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_r)^\top$  é um vetor de parâmetros lineares desconhecidos,  $\mathbf{x}[t] = (x_1[t], x_2[t], \dots, x_r[t])^\top$  é um vetor de variáveis determinísticas independentes e  $g: \mathbb{R}^+ \rightarrow \mathbb{R}$  é uma função de ligação estritamente monotônica e duas vezes diferenciável. Se o intercepto for considerado, então  $x_1[t] = 1$  para todo  $t$ . A função de ligação  $g(\cdot)$  relaciona os preditores lineares  $\eta[t] \in \mathbb{R}$  com os valores esperados  $\mu[t] \in \mathbb{R}^+$  de  $Y[t]$ . Como  $\mu[t] > 0$ , uma escolha comum para a função de ligação é a função log,  $\log(\mu[t]) = \eta[t]$ , com sua inversa  $\mu[t] = \exp(\eta[t])$ .

O modelo de regressão Rayleigh possui estrutura semelhante aos modelos lineares generalizados (GLMs) ([McCullagh and Nelder, 1989](#)), exceto pelo fato de a densidade Rayleigh não poder ser escrita na forma canônica da família exponencial de distribuições. Um modelo de regressão considerando a distribuição Rayleigh também é apresentado em [Aminzadeh \(1993\)](#). No entanto, este último é baseado na parametrização padrão da distribuição Rayleigh.

A partir desta proposta, alguns melhoramentos inferenciais para a regressão Rayleigh foram desenvolvidos. Em [Palm et al. \(2022b\)](#) foram estudadas correções de viés e suas aplicações em sinais de comprimento curto (pequenas amostras). O trabalho de [Palm et al. \(2022d\)](#) focou na proposta de estimadores robustos à presença de *outliers*. A abordagem é desenvolvida com base em estimadores de máxima verossimilhança ponderados. A distribuição Rayleigh inflacionada é proposta em [Palm et al. \(2022c\)](#), para contornar problemas de modelagem na presença de pixels nulos (falhas de sensor, etc.). Ou seja, esse trabalho generaliza a distribuição Rayleigh, propondo uma distribuição mista contínua-discreta para acomodar sinais que assumem valores em  $[0, \infty)$ . A respeito de modelos dinâmicos, a principal contribuição é o trabalho de [Bayer et al. \(2020b\)](#), o qual introduz o modelo RARMA. Esse modelo estende a regressão Rayleigh considerando um componente dinâmico adicional,  $\tau[t]$ , o qual acrescenta uma estrutura com termos autorregressivos (AR) e de médias móveis (MA) ao preditor linear.

Assim, seja  $\{Y[t]\}_{t \in \mathbb{Z}}$  um sinal a tempo discreto, em que cada  $Y[t]$  assume valores em  $\mathbb{R}^+$  e  $\mathcal{F}[t-1]$  é a sigma-álgebra gerada pelas informações até o tempo  $t-1$ . Assuma que, condicionalmente

a  $\mathcal{F}[t-1]$ , cada  $Y[t]$  é distribuído de acordo com uma distribuição Rayleigh parametrizada conforme (2.2), com média condicional  $\mu[t]$ .

A estrutura do modelo RARMA é dada pela Equação (2.3) e o componente adicional dinâmico,  $\tau[t]$ , da seguinte forma

$$g(\mu[t]) = \eta[t] = \mathbf{x}^\top[t]\boldsymbol{\beta} + \tau[t],$$

em que

$$\tau[t] = \sum_{i=1}^p \phi_i g(y[t-i]) + \sum_{j=1}^q \theta_j r[t-j],$$

onde  $\phi_i, i = 1, \dots, p$ , são os coeficientes AR,  $\theta_j, j = 1, \dots, q$ , são os coeficientes MA e  $r[t] = g(y[t]) - g(\mu[t])$  é o termo de erro na escala do preditor linear. Dessa forma, a estrutura dinâmica do modelo RARMA é definida por

$$g(\mu[t]) = \eta[t] = \zeta + \mathbf{x}^\top[t]\boldsymbol{\beta} + \sum_{i=1}^p \phi_i g(y[t-i]) + \sum_{j=1}^q \theta_j r[t-j],$$

em que  $\zeta \in \mathbb{R}$  é o intercepto. A média condicional de  $Y[t]$  é uma função das  $p$  observações passadas do sinal, covariáveis  $\mathbf{x}[t]$  e termos de erro MA. A função de ligação  $g(\cdot)$  conecta o preditor linear  $\eta[t]$  com a média condicional  $\mu[t]$ , da mesma forma que no modelo de regressão.

O modelo RARMA teve sua utilidade ilustrada mediante dados de velocidade do vento e imagens SAR multitemporais. Os resultados indicaram que esta estrutura representa uma alternativa flexível para modelar sinais a tempo discreto, assimétricos e não negativos, permitindo extrair características úteis para diferentes aplicações em sensoriamento remoto. Desta maneira, evidencia-se a relevância de se utilizar modelos dinâmicos com base na distribuição Rayleigh. É neste contexto que uma extensão do RARMA para caso bidimensional foi proposta e explorada por Palm et al. (2022a). Para facilitar a visualização da evolução cronológica dos avanços descritos nesta seção, estas contribuições são listadas na Tabela 2.1.

**Tabela 2.1:** Contribuições desenvolvidas a partir da distribuição Rayleigh reparametrizada na média.

Autor(es)	Modelo	Objetivo
Palm et al. (2019)	Regressão Rayleigh	Proposta do modelo
Bayer et al. (2020b)	RARMA	Inclui dinâmica ARMA ao modelo de regressão
Palm et al. (2022b)	Regressão Rayleigh	Correções de viés dos estimadores
Palm et al. (2022c)	Rayleigh inflacionado	Modelar sinais na presença de valores nulos
Palm et al. (2022d)	Regressão Rayleigh	Estimação robusta de parâmetros
Palm et al. (2022a)	2D RARMA	Considera o modelo bidimensional RARMA

### 2.3 A Especificação Básica dos Modelos GAS

Cox et al. (1981) classificaram os modelos com parâmetros variantes no tempo de duas maneiras: modelos guiados por observações e modelos guiados por parâmetros. Introduzida por Creal et al. (2008) e Creal et al. (2013), a classe GAS enquadra-se no primeiro caso, uma vez que são construídos com base em valores defasados da variável de interesse e do vetor escore ponderado da função de



verossimilhança. Seja  $\{Y[t]\}_{t \in \mathbb{Z}}$  uma sequência de variáveis aleatórias, cada uma com uma distribuição condicional dada por

$$Y[t] \sim f_Y(y[t] | f[t], \mathcal{F}[t-1], \theta), \quad t \in \mathbb{Z} \quad (2.4)$$

em que  $f[t]$  é um parâmetro variante no tempo,  $\theta$  é um vetor de parâmetros estáticos e  $\mathcal{F}[t-1] = \{y[1], \dots, y[t-1], x[1], \dots, x[t-1], f[1], \dots, f[t-1]\}$  é a sigma-álgebra gerada pelas informações até o tempo  $t-1$ . A especificação básica dos modelos GAS é dada pela seguinte equação de atualização

$$f[t] = \omega + \sum_{i=1}^p \alpha_i s[t-i] + \sum_{j=1}^q \beta_j f[t-j] + \sum_{k=1}^r \gamma_k x_k[t], \quad (2.5)$$

em que  $\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q$ , e  $\gamma_1, \dots, \gamma_r$  são componentes do vetor de parâmetros estáticos  $\theta$ ,  $x_k[t]$  é a  $k$ -ésima covariável no tempo  $t$ ,  $p, q$  e  $r$  são inteiros não negativos e  $s[t] = s(y[t], f[t]) = S[t] \cdot \nabla[t]$  é uma função dos valores passados que depende de  $S[t]$ , que é uma função de ponderação, e de  $\nabla[t]$ , que representa o elemento do vetor escore correspondente ao parâmetro variante no tempo. Temos que

$$\nabla[t] = \frac{\partial \log f_Y(y[t] | f[t], \mathcal{F}[t-1], \theta)}{\partial f[t]}. \quad (2.6)$$

Quando as observações do processo são realizadas até o instante  $t$ , o parâmetro variante no tempo é atualizado para o período  $t+1$  conforme a Equação (2.5). Assim, para uma dada escolha de  $S[t]$ , as expressões (2.4), (2.5) e (2.6) definem um modelo GAS de ordem  $p$  e  $q$ , denotado por  $\text{GAS}(p, q)$ . Segundo Creal et al. (2013), o uso de  $\nabla[t]$  para o mecanismo de atualização é intuitivo devido ao fato que, dada a posição de  $f[t]$ , a função escore indica a direção da subida mais acentuada para melhorar o ajuste local do modelo com base na verossimilhança, ou na densidade preditiva, no instante  $t$ . Os autores também observam que esta construção representa a estrutura completa da distribuição preditiva, pois a função escore depende da função de densidade na totalidade, e não apenas da média ou de momentos de ordem superior. Esta é a principal característica que difere os modelos GAS dos demais modelos guiados por observações explorados na literatura.

A escolha de  $S[t]$  pode ser feita de maneiras distintas. Entretanto, é natural considerá-la função da variância do vetor escore. Desta maneira, seja  $\mathcal{I}[t] = \mathbb{E}(\nabla[t] \nabla[t]^\top)$  a informação de Fisher da densidade preditiva considerada, as principais formas de ponderação encontradas na literatura são definidas como função do elemento de  $\mathcal{I}[t]$  correspondente ao parâmetro variante no tempo. De fato, nos trabalhos seminais de Creal et al. (2008) e Creal et al. (2013) são apresentadas possíveis escolhas para a função de ponderação que envolvem a informação de Fisher. Por exemplo, pode-se considerar  $S[t] = \mathcal{I}[t]^{-1}$ . Outras escolhas convenientes seriam  $S[t] = I$ , em que  $I$  é a matriz identidade, e  $S[t] = \mathcal{J}[t]$ , em que  $\mathcal{J}[t]^\top \mathcal{J}[t] = \mathcal{I}[t]^{-1}$  é a raiz quadrada da inversa da informação de Fisher para o elemento correspondente ao parâmetro variante no tempo.

Cabe ressaltar que a escolha de diferentes combinações de densidades preditivas e funções de ponderação implicam em um novo modelo GAS, com propriedades estatísticas distintas e potenciais aplicações diversas. Assim, diversas distribuições não Gaussianas podem ser utilizadas para definir novos modelos pertencentes à classe GAS. Além disso, ao permitir diferentes composições para  $S[t]$ , é permitida uma flexibilidade adicional na forma como a função escore é usada para atualizar  $f[t]$  (Creal et al., 2013). Esta flexibilidade é evidenciada ao observar que alguns modelos clássicos no contexto de séries temporais e finanças fazem parte da estrutura GAS. Além disso, contribuições sobre

a abordagem GAS têm sido cada vez mais exploradas na literatura de séries temporais e econometria. Esta informação pode ser verificada no site oficial dos modelos GAS ([Lucas and Koopman, 2023](#)), que reúne as principais informações sobre estes modelos e cataloga exaustivamente as contribuições acadêmicas da área. Além disso, na Tabela 2.2 são fornecidas informações acerca da especificação GAS para algumas densidades preditivas.

**Tabela 2.2:** Especificação GAS para alguns modelos.

Densidade	$f[t]$	$\nabla[t]$	$\mathcal{I}_{t/t-1}$
<b>Normal</b> $\frac{\exp[-(y[t]-\mu[t])^2/2\sigma[t]^2]}{\sqrt{2\pi\sigma[t]^2}}$	$\mu[t]$ $\sigma^2[t]$	$(y[t] - \mu[t]) / \sigma[t]^2$ $1/2\sigma^2 [-1 + (y[t] - \mu[t])^2 / \sigma[t]^2]$	$\mathcal{I}_{11}[t] = 1/\sigma[t]^2$ $\mathcal{I}_{22}[t] = 1/2\sigma[t]^4$ $\mathcal{I}_{12}[t] = 0$
<b>Exponencial</b> $\lambda[t] \exp(-\lambda[t])$	$\log(\lambda[t])$	$1 - y[t]\lambda[t]$	$\mathcal{I}_{11}[t] = 1$
<b>Gamma</b> $\frac{y[t]^{\alpha[t]-1} \exp(-y[t]/\beta[t])}{\Gamma(\alpha[t])\beta[t]^{\alpha[t]}}$	$\log(\alpha[t])$ $\log(\beta[t])$	$\alpha[t] [\log(y[t]) - \log(\beta[t]) + \psi_1(\alpha[t])]$ $y[t]/\beta[t] - \alpha[t]$	$\mathcal{I}_{11}[t] = \alpha[t]^2 \psi[2](\alpha[t])$ $\mathcal{I}_{22}[t] = \alpha[t]$ $\mathcal{I}_{12}[t] = \alpha[t]$
<b>Beta</b> $\frac{y[t]^{\alpha[t]-1} (1-y[t])^{\beta[t]-1}}{\beta(\alpha[t], \beta[t])}$	$\log(\alpha[t])$ $\log(\beta[t])$	$\alpha[t] [\psi[1](\alpha[t] + \beta[t]) - \psi[1](\alpha[t]) + \log(y[t])]$ $\beta[t] [\psi[1](\alpha[t] + \beta[t]) - \psi[1](\beta[t]) + \log(1 - y[t])]$	$\mathcal{I}_{11}[t] = \alpha[t]^2 [\psi[2](\alpha[t]) - \psi[2](\alpha[t] + \beta[t])]$ $\mathcal{I}_{22}[t] = \beta[t]^2 [\psi[2](\beta[t]) - \psi[2](\alpha[t] + \beta[t])]$ $\mathcal{I}_{12}[t] = -\alpha[t]\beta[t]\psi[2](\alpha[t] + \beta[t])$
<b>Poisson</b> $\exp(-\lambda[t]) \frac{\lambda[t]^{\psi[t]}}{\psi[t]!}$	$\log(\lambda[t])$	$(y[t] - \lambda[t])$	$\mathcal{I}_{11}[t] = \lambda[t]$
<b>Multinomial</b> $\frac{n! \prod_{j=1}^J p[j]^{\psi[j]}}{\prod_{j=1}^J \psi[j]!}$	$\log\left(\frac{p[it]}{1 - \sum_{j=1}^{j \neq i} p[j]}\right)$	$(y[it] - np[it])$	$\mathcal{I}_{ii}[t] = np[it](1 - p[it])$ $\mathcal{I}_{ij}[t] = -np[it]p[j]i$

Fonte: [Creal et al. \(2008\)](#)

---

## CAPÍTULO 3

# DESENVOLVIMENTO

---

As contribuições desta dissertação são expressas no formato de artigo científico. Desta maneira, o presente capítulo apresenta o texto completo do artigo intitulado “*The Rayleigh Generalized Autoregressive Score Model for SAR Data Interpretation*”, o qual representa o principal produto deste documento e será submetido a um periódico internacional.

No presente documento, é usada a notação  $f[t]$  ao invés de  $f_t$ , sendo esta última a notação comumente utilizada em séries temporais e estatística. Isso se deve ao fato de a notação indexada como  $f[t]$  ser a mais usual nas áreas de processamento de sinais/imagens, sendo o foco de nossa proposta, uma vez que propomos técnicas para modelagem estatística de imagens de sensoriamento remoto. O modelo proposto é ilustrado a partir de regiões de lago e floresta obtidos de uma imagem SAR que faz parte do sistema CARABAS II, as quais desempenham um papel fundamental na aplicação discutida no artigo. O Apêndice A deste documento fornece uma detalhada descrição da imagem utilizada, permitindo uma compreensão mais completa e detalhada das análises e conclusões apresentadas ao longo do texto principal.

# The Rayleigh Generalized Autoregressive Score Model for SAR Data Interpretation

Miguel Peña-Ramírez,<sup>\*</sup> Renata Rojas Guerra,<sup>†</sup> and Fábio M. Bayer<sup>‡</sup>

## Abstract

Amplitude returns of synthetic aperture radar (SAR) images present asymmetric behavior and strictly positive values, being adequately characterized by the Rayleigh distribution in homogeneous regions. In addition, the image pixels usually present some level of spatial dependence. This paper proposes the Rayleigh generalized autoregressive score (Ray-GAS) model, a dynamic model useful for SAR data interpretation. It is derived from the generalized autoregressive score (GAS) framework, assuming that the Rayleigh conditional mean is an index-varying parameter. We discuss conditional maximum likelihood estimation, diagnostic analysis, and prediction tools for the new GAS model. Additionally, we perform Monte Carlo simulations to evaluate the performance of the parameter estimators. Finally, the proposed methodology is applied to forest and lake region data of an image from the CARABAS II system. The results show the Ray-GAS model's benefit in understanding stochastic behavior and predicting SAR amplitude returns.

**Keywords**— GAS models, mean-based parametrization, Rayleigh distribution, remote sensing, SAR images.

## 1 Introduction

Synthetic aperture radar (SAR) systems are widely used in remote sensing, mainly due to their ability to operate day and night in adverse conditions and generate high-resolution images [1]. Their usefulness includes applications for resource exploration [2, 3], classification [4, 5], target detection [6, 7], disaster prevention [8, 9], and anomaly detection [10, 11]. In each situation, the statistical modeling of SAR images depends on several features, such as the geometry of the radar system used for data acquisition, the objective characteristics, and the formation of interference called speckle noise [12]. Hence, it is necessary to consider statistical models that adequately accommodate the behavior of a particular SAR system configuration [10]. In this context, the Rayleigh distribution is widely known as an appropriate descriptor of amplitude SAR data homogeneous regions [13]. This statistical model is considered due to its analytical tractability and ability to accommodate asymmetric behavior, especially in single-look SAR images.

Several studies have been performed in SAR imagery literature using Rayleigh-based methods to model amplitude returns. Most of these works assume that the parameters of the distribution are constant over the image and work under the independent and identically distributed (i.i.d.) supposition. That is the case of [14], which models SAR images with a Rayleigh distribution generalization. Other examples are [15] and [16], where the performance of the Rayleigh model is compared with other competitor statistical distributions. However, as emphasized by [11], the i.i.d. hypothesis is not always feasible since the image pixels may present a spatial dependence on resolution. Given this consideration, we can refer to the Rayleigh autoregressive moving average (RARMA) [2] model as a recent advance on the proposal of Rayleigh dynamical models for remotely sensed data. The RARMA model was introduced to fit Rayleigh-distributed discrete-time signals considering autoregressive and moving average (ARMA) terms, being useful for modeling signals with equispaced observations in time or space. Nevertheless, it is essential to point out that other dynamical structures can be developed to improve the interpretation of SAR image data.

---

<sup>\*</sup>M. Peña-Ramírez is with the Programa de Pós-graduação em Estatística, Universidade Federal de Rio Grande do Sul, Brazil (E-mail: [miguepe@gmail.com](mailto:miguepe@gmail.com)).

<sup>†</sup>R. R. Guerra is with the Departamento de Estatística, Universidade Federal de Santa Maria, Brazil and the Programa de Pós-graduação em Estatística, Universidade Federal de Rio Grande do Sul, Brazil (E-mail: [renata.r.guerra@ufsm.br](mailto:renata.r.guerra@ufsm.br)).

<sup>‡</sup>F. M. Bayer is with the Departamento de Estatística and LACESM, Universidade Federal de Santa Maria, Brazil, and the Programa de Pós-graduação em Estatística, Universidade Federal de Rio Grande do Sul, Brazil (E-mail: [bayer@ufsm.br](mailto:bayer@ufsm.br)).

The generalized autoregressive score (GAS) framework is an alternative dynamic model for analyzing non-Gaussian time series. It is a class of observation-driven models introduced by [17, 18]. The updating mechanism assumes that a time-varying parameter (or a vector of parameters) is a function of a scaled score of the conditional density and the past observations. The GAS framework is attractive since the scale score allows fully exploring the density structure to update the time-varying parameter(s) [18]. Besides, the development of extensions to accommodate asymmetry, long memory, and time-varying higher-order moments can be considered without introducing further complications [18]. For this reason, the primary motivation of this research is to introduce a GAS model for SAR data interpretation based on the Rayleigh conditional density, called the Rayleigh generalized autoregressive score (Ray-GAS) model.

The Ray-GAS model is pioneering in using the GAS class for modeling remote sensing images. For that, we assume that the conditional mean of the mean-based parameterized Rayleigh distribution serves as a time-varying parameter in the GAS framework. The advantage of this parametrization is its interpretation as a central tendency measure, being typically more helpful not only for obtaining dynamic structures but also in the regression context. We can refer the reader to [19–21] and [2] as contributions to the literature on non-Gaussian ARMA structures with mean-based parametrization. Despite this, there is much to address regarding the GAS class, especially for SAR image interpretation. Thus, the current work can contribute to filling this gap.

From [22], one can verify that the advantages of the GAS class are not explored in the remote sensing context. Most of the applications are focused on financial and econometric outcomes (see, for instance, [23, 24], and [25]). Accordingly, the second contribution of this work is to show the potential of the GAS class for SAR data interpretation. This goal is pursued by applying the novel dynamical model to amplitude returns extracted from rays of a measured SAR image. In this way, we aim to provide evidence that the Ray-GAS model can contribute to understanding stochastic behavior and predict remotely sensed signals, being an effective alternative to the RARMA model.

Specific techniques and algorithms are applied for image capture and processing in image processing. Images from the CARABAS II airborne system were collected to develop methods for detecting hidden targets in the forest [26]. Furthermore, CARABAS II images present areas with homogeneous areas. The data set is recently explored by [7], [11], [10], and [27]. The data from this system have also been considered to assess improved point estimators in the Rayleigh regression [28], evaluate the goodness-of-fit of clutter statistical models [15], and illustrate the usefulness of new SAR descriptors [29, 30]. This work focuses on modeling and prediction problems, showing the proposal’s applicability to analyze rays from forest and lake regions. The use of time series models for this purpose is a common practice. For example, [21] designed an ARMA model to describe the intensities of a SAR image ray. We can also refer to [31], [32], and [33], which apply time series models to squared windows or rays from image pixels. However, to the authors’ knowledge, this type of analysis has not been performed on images from the CARABAS II system.

The rest of the paper is organized as follows. Section 2 describes the mean-based Rayleigh distribution and the GAS framework. The proposed model is presented in Section 3, and its parameter estimation is in Section 4. We deliver the diagnostic and predictive tools used for the Ray-GAS model in Section 5. Monte Carlo simulations are performed, and Section 6.1 describes the results. Subsequently, we apply our model to actual SAR amplitude data from CARABAS II. This practical application is described in Section 6.2. Moreover, in Section 7, we draw the concluding remarks of the work.

## 2 Mathematical Setup

In this section, we briefly describe the properties of the Rayleigh distribution. We consider a mean-based parametrization, which has helped model remote sensing data sets. Subsequently, we describe the basic specification of the GAS models and some proposals framed in this framework.

### 2.1 The Mean-Based Rayleigh Distribution

In SAR image modeling, we are often interested in analyzing the amplitude returns of image pixels. The observed amplitude signals often exhibit non-Gaussian properties [34, 35], such as asymmetric distributions and strictly positive values. In this context, the Rayleigh distribution is widely used due to the excellent characterization of these properties.

In addition, its use is also justified by an essential physical interpretation. According to what is stated in [13, 14], and [36], when a signal pulse is reflected, the energy is scattered in various directions due to different factors. As the measured image is under a large number of returns, the central limit theorem can be invoked. According to this, the distributions of the real and imaginary parts of the complex signal are approximated Gaussian distributed, supposed to be independent, with zero mean and constant variance. Thus, the amplitude values of complex SAR data, i.e. the sum of squares of these components, are exactly Rayleigh distributed under these suppositions [37].

The Rayleigh mean-based parametrization was introduced by [7] to formulate the Rayleigh regression model. Under their proposal, a random variable  $Y$  has the Rayleigh distribution with the following density function

$$f_Y(y; \mu) = \frac{\pi y}{2\mu^2} \exp\left(-\frac{\pi y^2}{4\mu^2}\right), \quad (1)$$

where  $y > 0$  and  $\mu > 0$  is the mean parameter. The cumulative distribution function and the quantile function of  $Y$  are given by

$$F_Y(y; \mu) = 1 - \exp\left(-\frac{\pi y^2}{4\mu^2}\right), \quad (2)$$

and

$$Q_Y(u; \mu) = 2\mu \sqrt{\frac{-\log(1-u)}{\pi}}, \quad 0 < u < 1,$$

respectively. Under this reparametrization, the mean and variance are given, respectively, by

$$E(Y) = \mu,$$

and

$$\text{Var}(Y) = \mu^2 \left(\frac{4}{\pi} - 1\right).$$

Due to the straightforward interpretation of the mean, the above parametrization also has been used in other works related to remote sensing problems. For example, [28] studied bias-adjusted estimators for the Rayleigh regression parameter estimator and its applications at small signal lengths. [38] extended this model by employing a mixed continuous-discrete distribution for signals that assume values in  $[0, \infty)$ . [39] propose a Rayleigh regression model robust to the presence of outliers. Concerning time series analyses, the RARMA model [2] describes the mean of Rayleigh-distributed discrete-time signals using a dynamic structure that includes autoregressive (AR) and moving average (MA) terms. More recently, its two-dimensional extension was introduced by [11].

## 2.2 The Basic GAS Specification

The basic GAS specification takes the score vector to update the time-varying parameters. Let  $Y[1], \dots, Y[n]$  a sequence of random variables, each one with a conditional distribution given by

$$Y[t] \sim f_Y(y[t] | f[t], \mathcal{F}[t-1], \boldsymbol{\theta}), \quad t = 1, \dots, n,$$

where  $f[t]$  is the time-varying vector,  $\boldsymbol{\theta}$  is the static parameter vector, and  $\mathcal{F}[t-1] = \{y[1], \dots, y[t-1], f[1], \dots, f[t-1]\}$  is the available information up to time  $t$ .

In the GAS framework, the mechanism to update the time-varying parameter  $f[t]$  is given by the following dynamic structure

$$f[t] = \omega + \sum_{i=1}^p \alpha_i s[t-i] + \sum_{j=1}^q \beta_j f[t-j] + \sum_{k=1}^r \gamma_k x_k[t], \quad (3)$$

where  $\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q$  and  $\gamma_1, \dots, \gamma_r$  are components of the static parameter vector,  $x_k[t]$  is the value of  $k$ th covariate at time  $t$ ,  $p, q$ , and  $r$  are non-negative integers, and  $s[t] = s(y[t], f[t])$  is a function of past data. This way, when the time-varying parameter  $f[t]$  is observed in a time  $t+1$ , it is updated according to Equation (3).

The main characteristic of the GAS framework is to consider the score function to define  $s[t] = S[t]\nabla[t]$ , where

$$\nabla[t] = \frac{\partial \log f_Y(y[t] | f[t], \mathcal{F}[t-1], \boldsymbol{\theta})}{\partial f[t]},$$

is the score function corresponding to  $f[t]$ , and  $S[t]$  is a scaling factor usually taken as a function of the information matrix. We refer to [18] for some possible choices of the scaling factor. According to the authors, this feature of using  $\nabla[t]$  in the updating mechanism provides a natural direction for updating the GAS model parameters by capturing the most pronounced changes in the data series. By considering the differences between consecutive values, the model can better fit the series' evolution and provide more accurate results. In addition, as the score function depends on the complete density, the model introduces new data transformations to update the parameter  $f[t]$ .

The GAS models will differ according to the distribution, time-varying parameter, or updating mechanism in which it is framed. Thus, they can take on a wide variety of densities and allow different parametrization. Nonlinear models can also represent past information through the driving mechanism  $s[t]$ . Selecting the scaling factor  $S[t]$  allows flexibility in how the vector score is used to update  $f[t]$ .

We can refer to [40] for GAS models based on exponential-gamma, Weibull-gamma, and double-gamma mixture distributions among these different configurations. They were used to predict the volatility of financial time series. In addition, [41] proposed a three-parameter GAS model based on a zero-inflated negative binomial distribution. Recently, [42] suggested a GAS model based on the generalized bimodal Birnbaum-Saunders distribution. [43] introduces models based on Student's t-distribution, skewed generalized t-distribution, generalized exponential beta distribution of the second type, normal inverse Gaussian, and Meixner distributions. In all these examples, the authors emphasize that their models retain information about the density in each observation, being one of their main advantages.

### 3 The proposed Rayleigh-GAS Model

We now suggest the Ray-GAS model based on the GAS structure presented in the previous section. GAS models are used to model time series data, where  $f[t]$  is a time-varying parameter or a vector of time-varying parameters. However, from now on, we refer to  $f[t]$  as index-varying parameters. It is more general since we refer to pixels of the image, but without forgetting that the SAR image is built from the combination of multiple radar pulses emitted at different times and locations. Each radar pulse provides information about the amplitude of the reflected signal. Thus, the information collected by each line of pixels in a SAR image can be considered a time series that shows the evolution of the Earth's surface over time.

Let  $Y[1], \dots, Y[n]$ , be a signal of length  $n$  taking values in  $\mathbb{R}^+$ . We assume that each  $Y[t]$ ,  $t = 1, \dots, n$ , is conditionally Rayleigh distributed with conditional mean  $\mu[t]$ , under the parametrization described in Section 2.1. Thus, the conditional density function of  $Y[t]$ ,  $t = 1, \dots, n$ , in the Ray-GAS model is

$$f_Y(y[t] | f[t], \mathcal{F}[t-1]; \boldsymbol{\theta}) = \frac{\pi y[t]}{2\mu[t]^2} \exp\left(-\frac{\pi y[t]^2}{4\mu[t]^2}\right), \quad (4)$$

where  $\mathcal{F}[t-1]$  is a set of past information,  $f[t]$  is a time-varying parameter,  $y[t] > 0$  and  $\mu[t]$  is the index-varying parameter. From (4), it follows that the vector of the static parameters  $\boldsymbol{\theta} = (\omega, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q, \gamma_1, \dots, \gamma_r)^\top$  corresponds to the parameter vector inherent to the dynamical structure of the GAS model.

Since  $\mu[t]$  can only take positive values, we shall consider the log link function to define the Ray-GAS updating equation. Thus, for each new observation of the signal  $y[t]$ , the index-varying parameter  $\mu[t]$  is updated according to

$$\mu[t] = \exp(f[t]), \quad (5)$$

where  $f[t]$  satisfies (3). In the proposed model, the scaling matrix  $S[t]$  is given by the inverse of the expected Fisher information matrix corresponding to  $\mu[t]$ . Observing that

$$\nabla[t] = \frac{\pi y[t]^2}{2\mu[t]^2} - 2,$$

after some algebra, we get

$$s[t] = \frac{1}{4} \left( \frac{\pi y[t]^2}{2\mu[t]^2} - 2 \right).$$

Finally, we can explicitly write the updating equation for the index-varying mean parameter of the Ray-GAS as

$$\log(\mu[t]) = \omega + \sum_{i=1}^p \alpha_i \frac{\pi y[t-i] - 4\mu[t-i]^2}{8\mu[t-i]} + \sum_{j=1}^q \beta_j f[t-j] + \sum_{k=1}^r \gamma_k x_k[t]. \quad (6)$$

Thus, Equations (4) and (6) define the Ray-GAS model entirely.

### 4 Conditional Likelihood Inference

The parameter estimation described herein uses the conditional maximum likelihood theory [44]. The conditional maximum likelihood estimator (CMLEs) of  $\boldsymbol{\theta}$  is defined as

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \sum_{t=m+1}^n \ell[t],$$

where  $m = \max(p, q)$  and  $\ell[t]$  given by

$$\ell[t] = \log\left(\frac{\pi}{2}\right) + \log(y[t]) - \log(\mu[t]^2) - \frac{\pi y[t]^2}{4\mu[t]^2}, \quad (7)$$

with  $\mu[t]$  obtained from (6). In the rest of this section, we shall define the conditional score vector.



## 4.1 Conditional Score Vector

The conditional score vector comprises partial derivatives of the conditional log-likelihood function concerning each component of  $\boldsymbol{\theta}$ . Using the chain rule, one obtains

$$\mathbf{U}(\boldsymbol{\theta}) = \sum_{t=m+1}^n \frac{\partial \ell[t]}{\partial \boldsymbol{\theta}} = \sum_{t=m+1}^n \frac{\partial \ell[t]}{\partial \mu[t]} \frac{\partial \mu[t]}{\partial \boldsymbol{\theta}} \frac{\partial f[t]}{\partial f[\boldsymbol{\theta}]} = \sum_{t=m+1}^n \left( \frac{\pi y[t]^2}{2\mu[t]^2} - 2 \right) \frac{\partial f[t]}{\partial \boldsymbol{\theta}}.$$

The partial derivatives,  $\partial f[t]/\partial \boldsymbol{\theta}$ , are recursively computed as

$$\frac{\partial f[t]}{\partial \boldsymbol{\theta}} = \mathbf{z}[t] + \sum_{i=1}^p \alpha_i s[t-i]^{\mu[t-i]} \frac{\partial f[t-i]}{\partial \boldsymbol{\theta}} + \sum_{j=1}^q \beta_j \frac{\partial f[t-j]}{\partial \boldsymbol{\theta}}, \quad (8)$$

where  $\mathbf{z}[t] = (1, s[t-1], \dots, s[t-p], f[t-1], \dots, f[t-q], x_1[t], \dots, x_r[t])^\top$ ,

$$s[t-i]^{\mu[t-i]} = \frac{1}{g'(\mu[t-i])} \frac{\partial s[t-i]}{\partial \mu[t-i]} = \frac{-\pi y[t-i]^2}{4\mu[t-i]^2}.$$

and  $g'(\cdot)$  is the first derivative of the adopted link function  $g(\cdot)$ .

Let  $\mathbf{0}$  be a null vector of dimension  $\kappa = 1 + p + q + r$ . We must solve the system  $\mathbf{U}(\boldsymbol{\theta}) = \mathbf{0}$  to obtain the parameter estimates. We use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) optimization algorithm, which can converge faster than others and be a more straight and accurate option. The BFGS is an iterative quasi-Newton method for solving unconstrained nonlinear optimization problems and is one of the most widely used algorithms for numerical maximization.

## 5 Diagnostic Tools and Prediction

In this section, we introduce and describe some model selection criteria that can be used for model identification and present some diagnostic and prediction tools for the fitted Ray-GAS model. The diagnostic measures allow us to evaluate the quality-of-fit of the proposed model to an observed signal and compare it with other competitive models. A fitted model with adequate diagnostic measures can be used for out-of-sample prediction [45].

The information criteria help select the most suitable Ray-GAS model and also for comparison purposes. Akaike's (AIC) and Schwartz's (SIC) information criteria are the most popular in this context. They are obtained from the maximized conditional log-likelihood function. In this work, we use the modified versions of these criteria proposed by [46], not incorrectly, to penalize models with larger values of  $m$ . The modified AIC and SIC are given by

$$\text{MAIC}(\kappa) = -2\hat{\ell} \left( \frac{n}{n-m} \right) + 2\kappa$$

and

$$\text{MSIC}(\kappa) = -2\hat{\ell} \left( \frac{n}{n-m} \right) + \log(n)\kappa,$$

respectively, where  $\hat{\ell} = \sum_{t=m+1}^n \hat{\ell}[t]$  is the maximized conditional log-likelihood and  $\kappa$  is the number of model parameters.

The residuals of a fitted model are essential for evaluating the goodness-of-fit as they allow us to see how the observed and predicted values differ. Among the different types of residuals considered in the literature, we can refer to quantile residuals as an appropriate alternative for nonlinear and non-Gaussian dynamical models [47]. This type of residual is also widely used in GAS models, and some examples can be seen in [42, 47, 48]. As is defined in [49], the quantile residuals for the Ray-GAS are given by

$$r^{(q)}[t] = \Phi^{-1} \left( 1 - \exp \left( -\frac{\pi y[t]^2}{4\hat{\mu}[t]^2} \right) \right),$$

where  $\hat{\mu}[t]$  is the predicted value of  $\mu[t]$  obtained from Equation (10) and  $\Phi^{-1}$  denotes the quantile function of the standard normal distribution. Under the correct specification,  $r^{(q)}[t]$  follows the standard Gaussian distribution and presents no dependence over the indexes.

The Shapiro-Wilk (SW) test [50] can verify the null hypothesis that the quantile residuals follow the standard Gaussian distribution. The serial correlation and linear independence can be checked using the Ljung-Box (LB) test [51]. In the LB test, we are interested in testing

$$\begin{aligned} \mathcal{H}_0 &: \rho_1 = \rho_2 = \dots = \rho_l = 0 \\ \mathcal{H}_1 &: \text{at least one } \rho_k \neq 0, k = 1 \dots, l, \end{aligned}$$

where  $\rho_k = \text{cor}(r_t, r_{t+k})$  is the  $k$ th-order autocorrelation and  $l \in \mathbb{N}^+$ . Thus, the interest is testing the first  $l$  autocorrelations are zero. The LB test statistic is given by

$$Q_{\text{LB}} = n(n+2) \sum_{k=1}^l \frac{\hat{\rho}_k}{n-k}, \quad (9)$$

where  $\hat{\rho}_k$  is the sample autocorrelation at lag  $k$ ,  $k = 1, \dots, l$ . Under  $H_0$ , the test statistic  $Q_{\text{LB}}$  asymptotically follows a chi-square distribution with  $l - p - q$  degrees of freedom.

We perform in-sample and one-step-ahead prediction of the model, where each new observation in the series is updated with the actual value of the previous step. It is possible since all the inputs are known and under the assumption that the model is correctly fitted. Thus, for  $t = m + 1, \dots, n + 1$ , the predictions are obtained by sequentially setting

$$\hat{\mu}[t] = g^{-1} \left( \hat{\omega} + \sum_{i=1}^p \hat{\alpha}_i \frac{\pi y[t-i] - 4\hat{\mu}[t-i]^2}{8\hat{\mu}[t-i]} + \sum_{j=1}^q \hat{\beta}_j \hat{f}[t-j] + \sum_{k=1}^r x_k[t] \hat{\gamma}_k \right), \quad (10)$$

where  $\hat{f}[t]$  is obtained by replacing the parameters for their CMLEs in Equation (6). We set  $\hat{\mu}[t] = g^{-1}(\hat{\omega} + \sum_{k=1}^r x_k[t] \hat{\gamma}_k)$  for  $t = 1, \dots, m$ . One advantage of using the RMSE and MAPE is that they allow for comparing the prediction performance of models from different classes.

## 6 Numerical Experiments

In this section, we evaluate the performance of the proposed model by providing numerical experiments using simulated and measured SAR amplitude images. We illustrate the new dynamic model's suitability for SAR data interpretation and prediction, being a feasible alternative to the state-of-the-art RARMA model. For this purpose, we consider image patches related to forest and lake regions. The implementations and all the numerical experiments are carried out using the R programming language in version 4.2.2.2 [52]. The computer codes for fitting the Ray-GAS model are available upon request.

### 6.1 Analysis with Simulated Data

We perform Monte Carlo simulations to evaluate the performance of the CMLEs of the Ray-GAS parameter vector. The scenarios are defined from estimated parameters given by fitted Ray-GAS models for two patches of forest and lake regions selected from the measured SAR image in Fig. 2. For both regions. We select 300 lines with 100 pixels. Fit the Ray-GAS(1, 1) without covariates for each ray. We calculate the median of the CMLEs and take these values to define the parameter vectors used in the simulation experiment (see Fig. 3 for a graphical description of the parameter estimates in each ray). Hence, we have  $\theta = (-1.24, 0.76, 0.39)^\top$  in the forest scenario and  $\theta = (-1.73, 0.64, 0.24)^\top$  in the lake scenario. For each scenario, we compute 10,000 Monte Carlo replications and set the sizes of the simulated image rays at  $n \in \{100, 200, 500, 1000\}$ . We use the inversion method for data generation.

Table 1: Monte Carlo Simulation Results for Point estimation in Ray-GAS(1,1) model for Forest and Lake scenarios.

Scenario	Measure	$n$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\beta}_1$
Forest	RB%	100	12.358	-22.122	-19.393
		200	5.285	-2.149	-8.950
		500	1.694	-0.229	-3.515
		1000	0.658	-0.444	-2.158
	MSE	100	0.122	0.017	0.028
		200	0.049	0.012	0.016
		500	0.017	0.004	0.005
		1000	0.006	0.003	0.004
Lake	RB%	100	11.214	-3.677	-34.675
		200	5.726	-2.087	-18.119
		500	1.799	-0.375	-5.621
		1000	0.962	-0.188	-3.189
	MSE	100	0.173	0.015	0.031
		200	0.089	0.007	0.017
		500	0.023	0.002	0.004
		1000	0.01	0.001	0.002

To evaluate the performance of the CMLEs, we calculate the percent relative bias (RB%) and the mean squared error (MSE). We also compute the total relative bias (TRB) and total mean square error (TMSE) to provide a graphical analysis of the point estimators, corresponding to the sum of the absolute values of RB% and MSE, respectively. Table 1 presents the simulation results. As expected, both measures decrease when more extensive rays lengths are considered, agreeing with the consistency property of the CMLEs. It also observed a similar performance in both scenarios. For example, the  $\hat{\beta}_1$  presents the highest RB% for most sample sizes, followed by  $\hat{\omega}$ . The TRB and TMSE corroborate the appropriateness of the conditional maximum likelihood method for the Ray-GAS since the measures decay to zero as the sample size increases, see Fig. 1.

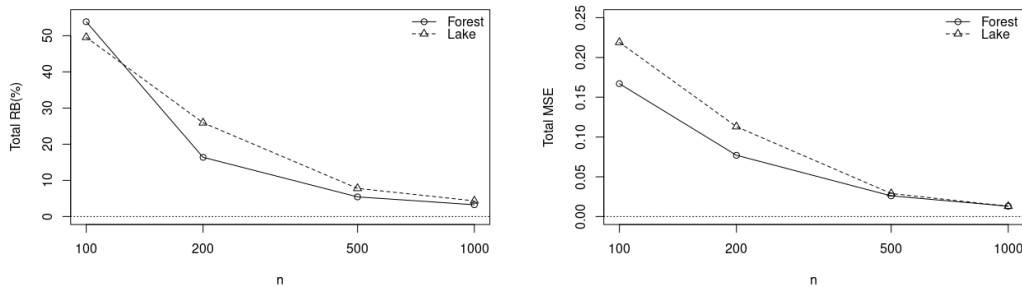


Figure 1: Total relative bias and mean square error of the model estimators in scenario forest and lake. (a) TRB (b) TMSE.

We also performed a numerical evaluation of the model selection and accuracy metrics when deciding between the Ray-GAS and RARMA models. We carry out the following procedure: i) Generate a signal of length  $n = 500$  under the Ray-GAS model with the parameter values from the forest scenario. ii) Fit the Ray-GAS and RARMA models. iii) Calculate both fitted models' MAIC, MSIC. iv) Verify which fitted model results in the smallest value for each measure. v) Replicate the experiment 1,000 times and, at the end of the replications, compute the percentage of times the measure selected the correct generation scheme. vi) Steps i to v are repeated for the lake scenario. vii) Steps i to vi are repeated for the RARMA model with the same parameter vectors as the Ray-GAS model. Table 2 reports the percentages of correct model selections achieved by the computed measures. They perform pretty well since all of their percentages lie above 99%. Therefore, we can conclude that these information criteria have reliable performance when comparing the Ray-GAS and RARMA data generation schemes.

Table 2: Percentage of correct model selection for the Ray-GAS and RARMA generation schemes.

Measure	Ray-GAS		RARMA	
	Forest	Lake	Forest	Lake
MAIC (%)	99.8	100.0	100.0	100.0
MSIC (%)	99.9	100.0	100.0	100.0

## 6.2 Analysis with Measured SAR Data

This section applies the Ray-GAS model to actual SAR amplitude data. The image of interest is extracted from the CARABAS II airborne system, provided by the Swedish Defense Research Agency and available at [53]. The system comprises ultrawideband very-high frequency SAR images from a forested region in northern Sweden and uses horizontal polarization. It mainly comprises lake and pine forest regions, and, as noted by [11], this area follows a homogeneous pattern and is a single-look image. The considered SAR image is shown in Fig. 2. It refers to a matrix of pixels with dimensions  $3000 \times 2000$  from the system's Mission 1. We select two patches that differ in their surface characteristics. As shown in Fig. 2, the first is located in a forest region, and the second corresponds to a lake region. Each patch is represented by a matrix of  $300 \times 125$  pixels comprising the positions  $[101, 400] \times [201, 325]$  and  $[2301, 2600] \times [1851, 1975]$  for the forest and lake regions, respectively. They are modeled by taking each line in the matrices as one ray of interest, i.e., the observed series is designed from the lines of the image patches. This modeling is joint for some imagery post-processing tasks [21]. For example, [32], [33], and [21] designed time series models based on the rays of the image pixels.

Our goal lies in comparing the Ray-GAS and RARMA models to identify the best class to model and predict the selected signals. To facilitate the interpretation and understanding of the model, we define  $p = q = 1$  for

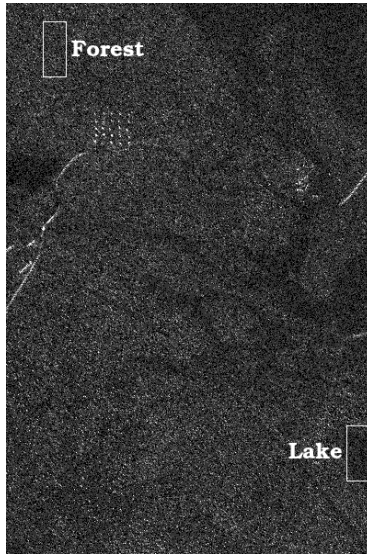


Figure 2: CARABAS II image used in the numerical experiments. The highlighted regions represent the forest and lake patches selected to fit the models.

fitting both dynamical classes. Also, we take the first  $n = 100$  (80%) pixels from each ray as the training set. We fit both competitive classes and compute the MAIC and MSIC using the training sets. Further, we checked the times that each class presented the best fit according to the calculated measures. Finally, the appropriateness of the quantile residuals is verified using the LB (with  $l = 20$ ) and SW tests at a significance level of 5%. The last  $n = 25$  (20%) pixels are used as a testing set to evaluate the one-step-ahead out-of-sample predictions. The predictions are obtained for each pixel in the test set by updating the series with the previous observation. After this, the RMSE and MAPE are computed considering the predictions in the testing set.

Fig. 3 provides boxplots that allow analyzing the behavior of the Ray-GAS estimated parameters. Notice that the lake region presents smaller estimates for the three model parameters, presenting lower median values than those from the forest. The lower median values in the lake region may indicate a lower backscatter in the same area [29]. Spatial information plays an important role, e.g., in class discrimination problems. Regarding the variability,  $\hat{\omega}$  has the highest interquartile range, and  $\hat{\alpha}_1$  provided the less dispersed values but with more outliers.

Table 3 presents the results regarding the LB and SW tests for the quantile residuals. Most of the rays return a  $p$ -value higher than 0.05 for both Ray-GAS and RARMA fitted models, indicating their residuals follow a normal distribution, and there is no evidence of dependence in the series. In the forest patch, the LB non-rejection rate is higher for the Ray-GAS model. When analyzing the lake region, both classes present a non-rejection rate of 81.33% for the LB and 73% for the SW tests. These diagnostic measures allow us to conclude that both classes are candidates for performing predictions based on the adjusted models.

Table 4 describes the results obtained for both training and testing sets analysis. All the measures indicate that the Ray-GAS model performs better in most rays. In the training set, the information criteria selected the proposed model in most rays, with much higher percentages than the RARMA model. The smallest percentage was 90.33% for the MAIC in the forest patch. The MAPE and RMSE ratios indicate that our proposal is slightly more accurate than the RARMA model in the testing set, with a higher percentage for the former. Therefore, the results from Tables 3 and 4 highlight the superiority of the Ray-GAS over the RARMA model when modeling and predicting the analyzed forest and lake patches. It illustrates the potential of the new dynamical model as an alternative for SAR data interpretation.

In order to explore a qualitative analysis, in what follows, we present a visual inspection to compare both fitted model classes in selected rays. Fig. 4 plots the observed and predicted values for each class in training set for a previously selected line from each region. We can observe that both models result in predicted values close to the observed signals, but the Ray-GAS model is more accurate, especially for the peaks of the series. Table 5 corroborates this information by providing the goodness-of-fit measures for these selected series. The Ray-GAS model outperformed the RARMA model in MAPE and RMSE by 2.056 and 0.003 percentage points, respectively, in the selected forest ray. In the lake series, it did so with 0.788 and 0.002 points in the same measures.

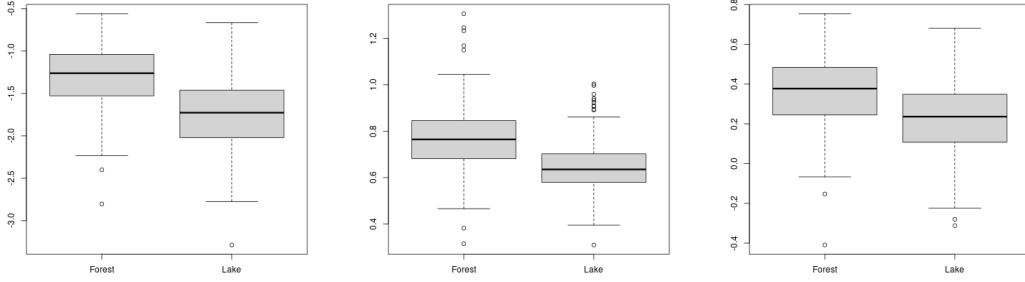


Figure 3: Boxplots of the Ray-GAS parameter estimates for both for each ray from the forest and lake patches. Values for (a)  $\hat{\omega}$ , (b)  $\hat{\alpha}_1$ , (c)  $\hat{\beta}_1$ .

Table 3: Number of times the null hypothesis is not rejected for the LB and SW tests in the training data sets. The percentages are in parentheses.

Patch	Model	LB	SW
Forest	<b>RARMA</b>	222 (74.00%)	217 (72.34%)
	<b>Ray-GAS</b>	230 (76.67%)	182 (60.67%)
	<b>RARMA and Ray-GAS</b>	192 (64.00%)	164 (54.67%)
Lake	<b>RARMA</b>	244 (81.33%)	258 (86.00%)
	<b>Ray-GAS</b>	244 (81.33%)	222 (74.00%)
	<b>RARMA and Ray-GAS</b>	225 (75.00%)	217 (72.33%)

Table 4: Number of times that the Ray-GAS model presented the best fit according to the selection criteria and accuracy measures. The percentages are in parentheses.

Measure/Criteria		Forest	Lake
Training set	<b>MAIC</b>	271 (90.33%)	285 (95.00%)
	<b>MSIC</b>	286 (95.33%)	294 (98.00%)
Testing set	<b>RMSE</b>	160 (53.33%)	157 (52.33%)
	<b>MAPE</b>	160 (53.33%)	157 (52.33%)

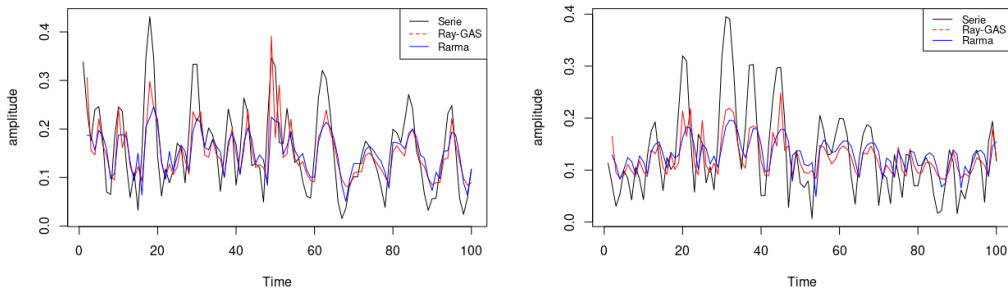


Figure 4: Observed and predicted values for the training set of rays selected from the (a) forest, and (b) lake patches.

Table 5: Accuracy measures for the Ray-GAS and RARMA models fitted in rays selected from the forest and lake patches.

	Forest		Lake	
	MAPE	RMSE	MAPE	RMSE
<b>Ray-GAS</b>	35.990	0.056	33.923	0.044
<b>RARMA</b>	38.046	0.059	34.711	0.046

Regarding the residuals' analysis, the LB, and SW tests, with a significance level of  $\alpha = 0.05$ , suggest that the quantile residuals are independent and Gaussian-distributed in all cases. This conclusion is confirmed by their residual autocorrelation function (ACF) and the partial autocorrelation function (PACF). These plots are shown in Fig. 5 and Fig. 6 for the forest and lake region, respectively.

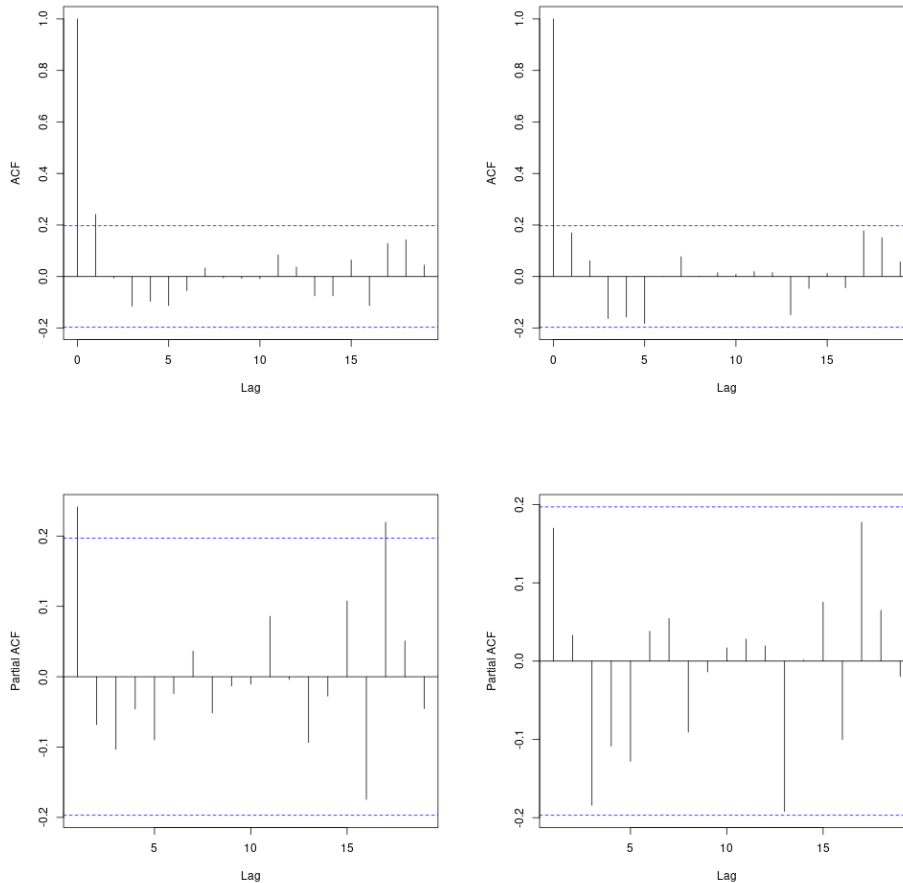


Figure 5: Autocorrelation functions for the residuals of the training set of the ray selected from the forest patch. (a) ACF RARMA, (b) ACF Ray-GAS, (c) PACF RARMA, (d) PACF Ray-GAS.

## 7 Concluding Remarks

This paper proposed a dynamical model based on the GAS structure and the Rayleigh conditional distribution to interpret SAR data. The parameter estimation was performed using the conditional maximum likelihood method, and some diagnostic tools were suggested to evaluate the quality-of-fit of the new model. We also presented the procedure to obtain one-step-ahead predictions. A Monte Carlo simulation study was carried out to assess the performance of the proposed estimators and the model selection criteria. These simulations showed that the suggested estimation method suits the Ray-GAS model. Therefore, the information criteria presented a reliable performance when comparing the Ray-GAS and RARMA data generation schemes. As applications to measured SAR data, we have employed our model to describe amplitudes of forest and lake regions from a CARABAS II SAR image. Results have shown that the proposed model can perform better than the RARMA model, being an effective alternative to this model. These findings reveal that the proposed model can improve the understanding of stochastic behavior and predict SAR amplitude returns, evidencing the potential of the GAS class for SAR data interpretation.

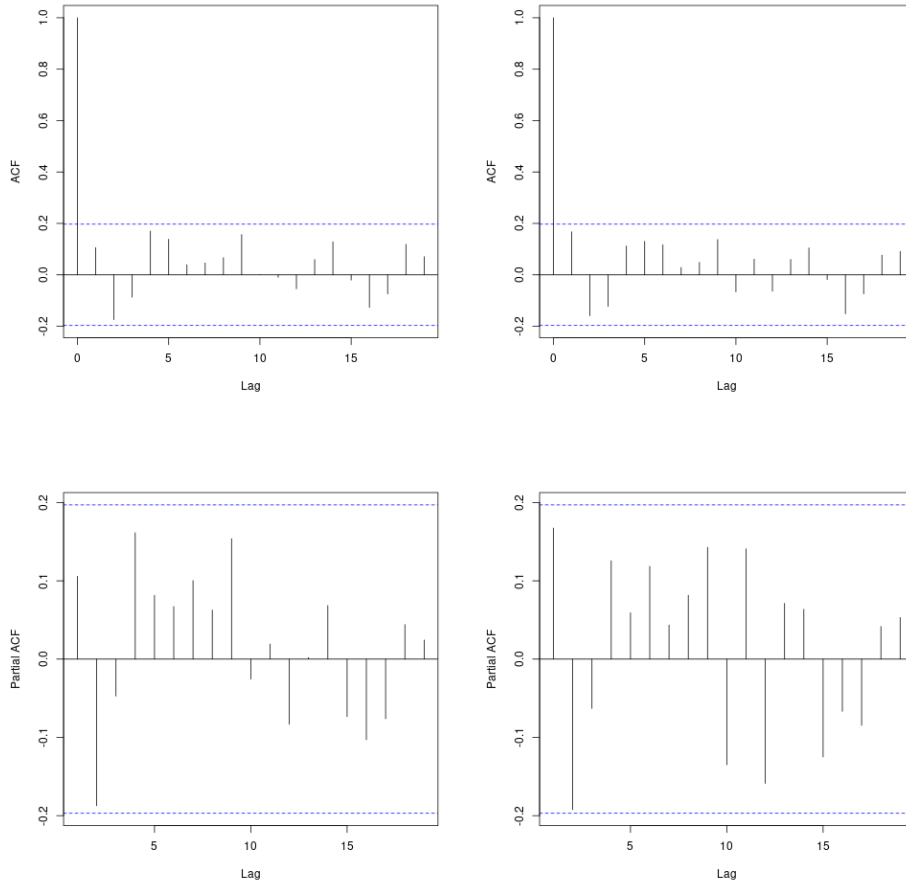


Figure 6: Autocorrelation functions for the residuals of the training set of the ray selected from the lake patch. (a) ACF RARMA, (b) ACF Ray-GAS, (c) PACF RARMA, (d) PACF Ray-GAS.

## References

- [1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, “A tutorial on synthetic aperture radar,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, pp. 6–43, 2013.
- [2] F. M. Bayer, D. M. Bayer, A. Marinoni, and P. Gamba, “A novel Rayleigh dynamical model for remote sensing data interpretation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, pp. 4989–4999, 2020.
- [3] R. Avtar, N. Sahu, A. K. Aggarwal, S. Chakraborty, A. Kharrazi, A. P. Yunus, J. Dou, and T. A. Kurniawan, “Exploring renewable energy resources using remote sensing and GIS—A review,” *Resources*, vol. 8, no. 3, p. 149, 2019.
- [4] H. Parikh, S. Patel, and V. Patel, “Classification of SAR and polSAR images using deep learning: A review,” *International Journal of Image and Data Fusion*, vol. 11, no. 1, pp. 1–32, 2020.
- [5] H.-W. Jo, S. Lee, E. Park, C.-H. Lim, C. Song, H. Lee, Y. Ko, S. Cha, H. Yoon, and W.-K. Lee, “Deep learning applications on multitemporal SAR (Sentinel-1) image classification using confined labeled data: The case of detecting rice paddy in South Korea,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 11, pp. 7589–7601, 2020.
- [6] H. Lang, Y. Xi, and X. Zhang, “Ship detection in high-resolution SAR images by clustering spatially enhanced pixel descriptor,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 8, pp. 5407–5423, 2019.
- [7] B. G. Palm, F. M. Bayer, R. J. Cintra, M. I. Pettersson, and R. Machado, “Rayleigh regression model for ground type detection in SAR imagery,” *IEEE Geoscience and Remote Sensing Letters*, vol. 16, pp. 1660–1664, 2019.

- [8] M. Turgeon-Pelchat, H. McGrath, F. Esfahani, S. Tolszczuk-Leclerc, T. Rainville, N. Svacina, L. Zhou, Z. Langari, and H. Houngho, "Improving near real-time flood extraction pipeline from SAR data using deep learning," tech. rep., Copernicus Meetings, 2023.
- [9] D. Song, Z. Zhen, B. Wang, X. Li, L. Gao, N. Wang, T. Xie, and T. Zhang, "A novel marine oil spillage identification scheme based on convolution neural network feature extraction from fully polarimetric SAR imagery," *IEEE Access*, vol. 8, pp. 59801–59820, 2020.
- [10] M. Sagrillo, R. Guerra, R. Machado, and F. Bayer, "A generalized control chart for anomaly detection in SAR imagery," *Computers & Industrial Engineering*, p. 109030, 2023.
- [11] B. G. Palm, F. M. Bayer, and R. J. Cintra, "2-D Rayleigh autoregressive moving average model for SAR image modeling," *Computational Statistics & Data Analysis*, vol. 171, p. 107453, 2022.
- [12] E. Podest, "Basics of synthetic aperture radar," *NASA Applied Remote Sensing Training (ARSET) Program*, 2017.
- [13] C. Oliver and S. Quegan, *Understanding synthetic aperture radar images*. SciTech Publishing, 2004.
- [14] E. E. Kuruoglu and J. Zerubia, "Modeling SAR images with a generalization of the Rayleigh distribution," *IEEE Transactions on Image Processing*, vol. 13, pp. 527–533, 2004.
- [15] D. I. Alves, B. G. Palm, M. I. Pettersson, V. T. Vu, R. Machado, B. F. Uchoa-Filho, P. Dammert, and H. Hellsten, "A statistical analysis for wavelength-resolution SAR image stacks," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, pp. 227–231, 2019.
- [16] D.-X. Yue, F. Xu, A. Frery, and Y.-Q. Jin, "A generalized Gaussian coherent scatterer model for correlated SAR texture," -, 2020.
- [17] D. Creal, S. J. Koopman, and A. Lucas, "A general framework for observation driven time-varying parameter models," -, 2008.
- [18] D. Creal, S. J. Koopman, and A. Lucas, "Generalized autoregressive score models with applications," *Journal of Applied Econometrics*, vol. 28, pp. 777–795, 2013.
- [19] M. A. Benjamin, R. A. Rigby, and D. M. Stasinopoulos, "Generalized autoregressive moving average models," *Journal of the American Statistical Association*, vol. 98, no. 461, pp. 214–223, 2003.
- [20] A. V. Rocha and F. Cribari-Neto, "Beta autoregressive moving average models," *Test*, vol. 18, no. 3, p. 529, 2009.
- [21] P. M. Almeida-Junior and A. D. Nascimento, "ARMA process for speckled data," *Journal of Statistical Computation and Simulation*, vol. 91, pp. 3125–3153, 2021.
- [22] A. Lucas, "Generalized autoregressive score models," 2008. [Online] Available: <http://www.gasmodel.com/gaspapers.htm> [Web; Accessed on 03-02-2023].
- [23] W. Alanya-Beltran, "Modelling stock returns volatility with dynamic conditional score models and random shifts," *Finance Research Letters*, vol. 45, p. 102121, 2022.
- [24] F. Fuentes, R. Herrera, and A. Clements, "Forecasting extreme financial risk: A score-driven approach," *International Journal of Forecasting*, 2022.
- [25] M. Folqué, E. Escrig-Olmedo, and T. Corzo Santamaría, "Sustainable development and financial system: Integrating esg risks through sustainable investment strategies in a climate change context," *Sustainable Development*, vol. 29, no. 5, pp. 876–890, 2021.
- [26] M. Lundberg, L. M. H. Ulander, W. E. Pierson, and A. Gustavsson, "A challenge problem for detection of targets in foliage," in *Algorithms for Synthetic Aperture Radar Imagery XIII*, vol. 6237, p. 62370K, SPIE, 2006.
- [27] L. P. Ramos, A. B. Campos, C. Schwartz, L. T. Duarte, D. I. Alves, M. I. Pettersson, V. T. Vu, and R. Machado, "A wavelength-resolution SAR change detection method based on image stack through robust principal component analysis," *Remote Sensing*, vol. 13, p. 833, 2021.
- [28] B. G. Palm, F. M. Bayer, and R. J. Cintra, "Improved point estimation for the Rayleigh regression model," *IEEE Geoscience and Remote Sensing Letters*, 2022.
- [29] M. Sagrillo, R. R. Guerra, F. M. Bayer, and R. Machado, "A new probability distribution for SAR image modeling," *Remote Sensing*, vol. 14, p. 2853, 2022.
- [30] M. Sagrillo, R. R. Guerra, F. M. Bayer, and R. Machado, "The  $\mathcal{L}_a$  distribution: An approximation of the  $\mathcal{G}_a^0$  distribution for amplitude SAR image modeling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. pre-print, 2023.
- [31] E. J. Delp, R. L. Kashyap, and O. R. Mitchell, "Image data compression using autoregressive time series models," *Pattern Recognition*, vol. 11, pp. 313–323, 1979.



- [32] R. T. Chin, J. Y. Jau, and J. A. Weinman, "The application of time series models to cloud field morphology analysis," *Journal of Applied Meteorology and Climatology*, vol. 26, pp. 363–373, 1987.
- [33] R. M. Rangayyan, *Biomedical image analysis*. CRC press, 2004.
- [34] A. Margosian, J. Abouei, and K. N. Plataniotis, "An accurate kernelized energy detection in Gaussian and non-Gaussian/impulsive noises," *IEEE Transactions on Signal Processing*, vol. 63, no. 21, pp. 5621–5636, 2015.
- [35] J. Liu, S. Liu, W. Liu, S. Zhou, S. Zhu, and Z.-J. Zhang, "Persymmetric adaptive detection of distributed targets in compound-Gaussian sea clutter with gamma texture," *Signal Processing*, vol. 152, pp. 340–349, 2018.
- [36] D.-X. Yue, F. Xu, A. C. Frery, and Y.-Q. Jin, "Synthetic aperture radar image statistical modeling: Part one-single-pixel statistical models," *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, pp. 82–114, 2020.
- [37] C. Oliver, "Information from SAR images," *Journal of Physics D: Applied Physics*, vol. 24, no. 9, p. 1493, 1991.
- [38] B. G. Palm, S. Javadi, F. M. Bayer, V. T. Vu, and M. I. Pettersson, "Inflated Rayleigh distribution for SAR imagery modeling," in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*, pp. 44–47, IEEE, 2022.
- [39] B. G. Palm, F. M. Bayer, R. Machado, M. I. Pettersson, V. T. Vu, and R. J. Cintra, "Robust Rayleigh regression method for SAR image processing in presence of outliers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022.
- [40] S. J. Koopman, A. Lucas, and M. Scharth, "Predicting time-varying parameters with parameter-driven and observation-driven models," *Review of Economics and Statistics*, vol. 98, pp. 97–110, 2016.
- [41] F. Blasques, V. Holý, and P. Tomanová, "Zero-inflated autoregressive conditional duration model for discrete trade durations with excessive zeros," *arXiv preprint arXiv:1812.07318*, 2018.
- [42] R. V. Fonseca and F. Cribari-Neto, "Bimodal Birnbaum–Saunders generalized autoregressive score model," *Journal of Applied Statistics*, vol. 45, pp. 2585–2606, 2018.
- [43] A. Ayala, S. Blazsek, and A. Licht, "Score-driven stochastic seasonality of the Russian rouble: an application case study for the period of 1999 to 2020," *Empirical Economics*, vol. 62, no. 5, pp. 2179–2203, 2022.
- [44] E. B. Andersen, "Asymptotic properties of conditional maximum-likelihood estimators," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 32, no. 2, pp. 283–301, 1970.
- [45] W. K. Li, *Diagnostic checks in time series*. Chapman and Hall/CRC, 2003.
- [46] F. M. Bayer, R. J. Cintra, and F. Cribari-Neto, "Beta seasonal autoregressive moving average models," *Journal of Statistical Computation and Simulation*, vol. 88, pp. 2961–2981, 2018.
- [47] H. Hoeltgebaum, C. Fernandes, and A. Street, "Generating joint scenarios for renewable generation: The case for non-Gaussian models with time-varying parameters," *IEEE Transactions on Power Systems*, vol. 33, pp. 7011–7019, 2018.
- [48] M. A. B. de Melo, C. A. Fernandes, and E. F. de Melo, "Forecasting aggregate claims using score-driven time series models," *Statistica Neerlandica*, vol. 72, pp. 354–374, 2018.
- [49] P. K. Dunn and G. K. Smyth, "Randomized quantile residuals," *Journal of Computational and graphical statistics*, vol. 5, pp. 236–244, 1996.
- [50] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965.
- [51] G. M. Ljung and G. E. Box, "On a measure of lack of fit in time series models," *Biometrika*, vol. 65, pp. 297–303, 1978.
- [52] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2022.
- [53] SDMS, "Sensor data management system," 2002. [Online] Available: [www.site.com.br/index.html](http://www.site.com.br/index.html) [Web; Accessed on 01-2023].

---

## CAPÍTULO 4

# CONCLUSÕES E TRABALHOS FUTUROS

---

Nesta dissertação foi proposto um modelo dinâmico para modelagem e previsão de imagens de radar de abertura sintética (SAR). O novo modelo é introduzido a partir da classe de modelos escore autorregressivo generalizado (GAS), assumindo a distribuição Rayleigh como densidade preditiva. A classe GAS se diferencia dos demais modelos dinâmicos por utilizar informações da verossimilhança na atualização dos parâmetros variantes no tempo. Os tradicionais modelos autorregressivos de médias móveis, por exemplo, não apresentam essa vantagem, uma vez que consideram apenas a média ou algum quantil da variável resposta em seu mecanismo de atualização. Outra vantagem da estrutura GAS é a facilidade para acomodar uma ampla variedade de densidades preditivas e suas diferentes parametrizações, tornando esta classe uma importante alternativa para analisar séries temporais não Gaussianas. Assim, uma vez que os retornos em amplitude de imagens SAR apresentam comportamento assimétrico e valores estritamente positivos, modelos GAS construídos com base em distribuições que acomodam este comportamento se mostram boas alternativas a serem aplicadas neste tipo de problema prático. É nesse contexto que se inserem as contribuições desta dissertação, uma vez que a distribuição Rayleigh atende a estas características e é um modelo amplamente explorado não apenas em aplicações de imagem SAR, mas também em outros problemas de sensoriamento remoto de maneira geral.

Além da proposição do modelo, seus aspectos inferenciais são abordados através do método de máxima verossimilhança condicional. O vetor escore condicional é obtido em forma fechada. Também foi apresentado um conjunto de ferramentas para análise de diagnóstico do modelo, tais como métricas para a seleção de modelos e avaliação da qualidade do ajuste. Além disso, é sugerido o uso dos resíduos quantílicos em conjunto com a realização de testes de hipóteses para a verificação da adequabilidade do modelo ajustado. A expressão para obtenção de previsões um passo à frente também foi apresentada. Na estrutura GAS, estes valores preditos podem ser obtidos diretamente da distribuição condicional de observações até o tempo atual.

O desempenho do novo modelo foi avaliado por meio de experimentos numéricos com imagens SAR mensuradas e simuladas. O intuito das simulações foi avaliar os estimadores de máxima verossimilhança condicional e verificar a adequabilidade das métricas propostas para a seleção de modelos, comparando os modelos Rayleigh autorregressivos de médias móveis e Ray-GAS. O estudo de simulação evidenciou que o método de estimação pontual utilizado é apropriado para o modelo Ray-GAS e que as métricas consideradas conseguem distinguir o processo gerador dos dados ao comparar as classes de modelos consideradas. Estes resultados viabilizam o uso do modelo Ray-GAS como uma alternativa ao modelo RARMA para interpretação e previsão de dados SAR. Assim, seguiu-se com a comparação

destes modelos em imagens mensuradas do sistema CARABAS II. A imagem SAR de interesse é analisada com foco em problemas de modelagem e predição. Os modelos RARMA e Ray-GAS foram ajustados em diferentes linhas de imagens SAR provenientes de regiões de floresta e lago. Os resultados indicam a superioridade da proposta para a maioria das linhas analisadas, confirmando que o modelo Ray-GAS pode proporcionar um melhor desempenho que o RARMA na interpretação de imagens SAR.

O presente trabalho pode abrir várias áreas de atuação no que tange modelos GAS para modelagem e interpretação de imagens. Pontualmente, destacamos alguns avanços que podem ser desenvolvidos como trabalhos futuros:

- Estudo da aplicabilidade do novo modelo em outros problemas de sensoriamento remoto, como é o caso na velocidade do vento e de dados multitemporais, em que se obtém uma série temporal univariada para cada píxel. Este tipo de análise foi realizada por [Bayer et al. \(2020b\)](#) utilizando o modelo RARMA. É esperado que o modelo proposto tenha potencial para estas aplicações, uma vez que se apresentou competitivo ao RARMA no estudo empírico realizado nesta dissertação.
- Desenvolvimento e disponibilização de um pacote computacional com as implementações relacionadas ao modelo proposto, viabilizando o uso do modelo por outros pesquisadores.
- Desenvolvimento de extensões do modelo Ray-GAS como modelos bidimensionais que considerem outras vizinhanças do píxel de interesse. Uma abordagem deste tipo é considerada em [Palm et al. \(2022a\)](#) como uma extensão espacial do modelo RARMA unidimensional. Apesar de haver alguns avanços em acomodar comportamentos espaço-temporais na estrutura GAS, esta abordagem relacionada à modelagem de imagens digitais ainda não foi introduzida para os modelos GAS.

---

## REFERÊNCIAS BIBLIOGRÁFICAS

---

- Alanya-Beltran, W., 2022. Modelling stock returns volatility with dynamic conditional score models and random shifts. *Finance Research Letters* 45, 102121.
- Almeida-Junior, P.M., Nascimento, A.D., 2021. ARMA process for speckled data. *Journal of Statistical Computation and Simulation* 91, 3125–3153.
- Aminzadeh, M., 1993. Approximate 1-sided tolerance limits for future observations for the Rayleigh distribution, using regression. *IEEE Transactions on Reliability* 42, 625–630.
- Bayer, D.M., Bayer, F.M., Gamba, P., 2020a. A 3-D spatiotemporal model for remote sensing data cubes. *IEEE Transactions on Geoscience and Remote Sensing* 59, 1082–1093.
- Bayer, F.M., Bayer, D.M., Marinoni, A., Gamba, P., 2020b. A novel Rayleigh dynamical model for remote sensing data interpretation. *IEEE Transactions on Geoscience and Remote Sensing* 58, 4989–4999.
- Chen, J., Wang, Q., Wang, J., Li, N., 2019. Change detection of water index in Danjiangkou reservoir using mixed log-normal distribution based active contour model. *IEEE Access* 7, 95430–95442.
- Cintra, R.J., Rêgo, L.C., Cordeiro, G.M., Nascimento, A.D., 2012. Beta generalized normal distribution with an application for SAR image processing. *Statistics* 48, 279–294.
- Cordeiro, G.M., Cintra, R.J., Rego, L.C., Nascimento, A.D., 2019. The gamma generalized normal distribution: A descriptor of SAR imagery. *Journal of Computational and Applied Mathematics* 347, 257–272.
- Cox, D.R., Gudmundsson, G., Lindgren, G., Bondesson, L., Harsaae, E., Laake, P., Juselius, K., Lauritzen, S.L., 1981. Statistical analysis of time series: Some recent developments [with discussion and reply]. *Scandinavian Journal of Statistics* , 93–115.
- Creal, D., Koopman, S.J., Lucas, A., 2008. A general framework for observation driven time-varying parameter models. *Tinbergen Institute Discussion Paper* 08-108/4.
- Creal, D., Koopman, S.J., Lucas, A., 2013. Generalized autoregressive score models with applications. *Journal of Applied Econometrics* 28, 777–795.
- Eineder, M., Minet, C., Steigenberger, P., Cong, X., Fritz, T., 2010. Imaging geodesy—toward centimeter-level ranging accuracy with TerraSAR-X. *IEEE Transactions on Geoscience and Remote Sensing* 49, 661–671.
- Engle, R., 2002. New frontiers for ARCH models. *Journal of Applied Econometrics* 17, 425–446.

- Engle, R.F., Bollerslev, T., 1986. Modelling the persistence of conditional variances. *Econometric Reviews* 5, 1–50.
- Engle, R.F., Russell, J.R., 1998. Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica* , 1127–1162.
- Folqué, M., Escrig-Olmedo, E., Corzo Santamaría, T., 2021. Sustainable development and financial system: Integrating esg risks through sustainable investment strategies in a climate change context. *Sustainable Development* 29, 876–890.
- Fuentes, F., Herrera, R., Clements, A., 2023. Forecasting extreme financial risk: A score-driven approach. *International Journal of Forecasting* 39, 20–735.
- Gao, G., 2010. Statistical modeling of SAR images: A survey. *Sensors* 10, 775–795.
- Kuruoglu, E.E., Zerubia, J., 2004. Modeling SAR images with a generalization of the Rayleigh distribution. *IEEE Transactions on Image Processing* 13, 527–533.
- Lee, J.S., Pottier, E., 2017. *Polarimetric radar imaging: from basics to applications*. CRC press.
- Li, H.C., Hong, W., Wu, Y.R., Fan, P.Z., 2010. An efficient and flexible statistical model based on generalized gamma distribution for amplitude SAR images. *IEEE Transactions on Geoscience and Remote Sensing* 48, 2711–2722.
- Liu, X., He, C., Zhang, Q., Liao, M., 2019. Statistical convolutional neural network for land-cover classification from SAR images. *IEEE Geoscience and Remote Sensing Letters* 17, 1548–1552.
- Lucas, A., Koopman, S.J., 2023. Generalized autoregressive score models. Available: <http://www.gasmodel.com/gaspapers.htm> [Web; Accessed on 03-02-2023].
- McCullagh, P., Nelder, J., 1989. *Generalized linear models*. 2 ed., Chapman & Hall, London, UK.
- Meng, Q., Wen, X., Yuan, L., Xu, H., 2019. Factorization-based active contour for water-land SAR image segmentation via the fusion of features. *IEEE Access* 7, 40347–40358.
- Merkle, N., Auer, S., Müller, R., Reinartz, P., 2018. Exploring the potential of conditional adversarial networks for optical and SAR image matching. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11, 1811–1820.
- Meyer, F.J., Chotoo, K., Chotoo, S.D., Huxtable, B.D., Carrano, C.S., 2016. The influence of equatorial scintillation on L-Band SAR image quality and phase. *IEEE Transactions on Geoscience and Remote Sensing* 54, 869–880.
- Moreira, M.A., 2005. *Fundamentos do sensoriamento remoto e metodologias de aplicação*. UFV.
- Oliver, C., Quegan, S., 2004. *Understanding synthetic aperture radar images*. SciTech Publishing.
- Palm, B.G., Bayer, F.M., Cintra, R.J., 2022a. 2-D Rayleigh autoregressive moving average model for SAR image modeling. *Computational Statistics & Data Analysis* 171, 107453.
- Palm, B.G., Bayer, F.M., Cintra, R.J., Pettersson, M.I., Machado, R., 2019. Rayleigh regression model for ground type detection in SAR imagery. *IEEE Geoscience and Remote Sensing Letters* 16, 1660–1664.

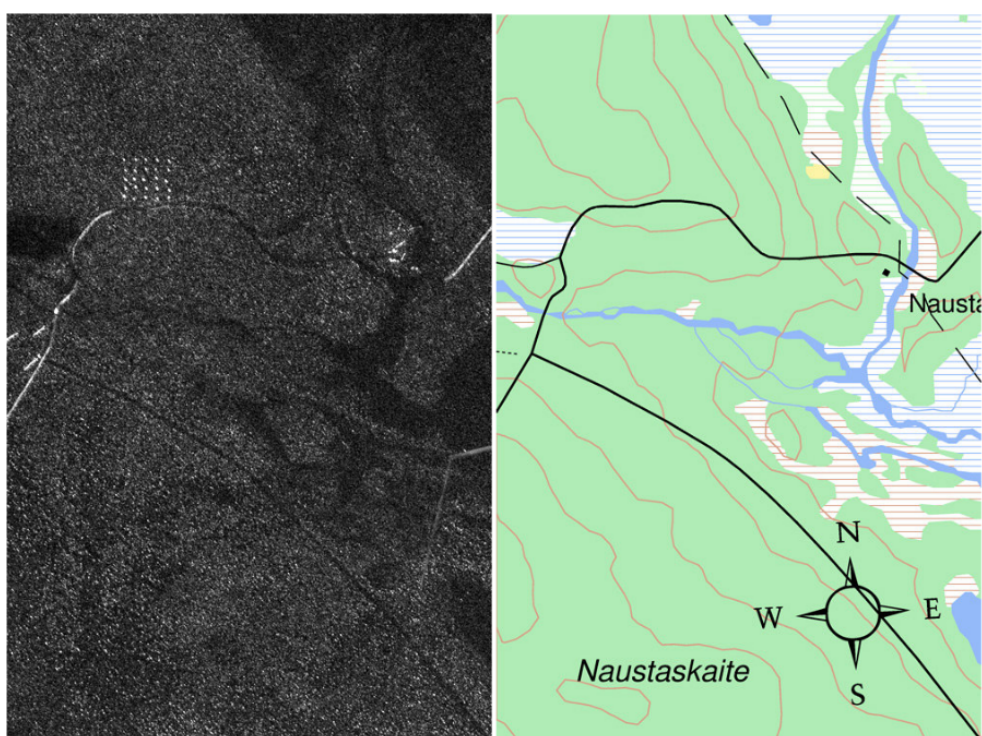
- Palm, B.G., et al., 2022b. Improved point estimation for the Rayleigh regression model. *IEEE Geoscience and Remote Sensing Letters* 19, 1–4.
- Palm, B.G., et al., 2022c. Inflated Rayleigh distribution for SAR imagery modeling, in: *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*, IEEE. pp. 44–47.
- Palm, B.G., et al., 2022d. Robust Rayleigh regression method for SAR image processing in presence of outliers. *IEEE Transactions on Geoscience and Remote Sensing* 60, 1–12.
- Peng, Q., Zhao, L., 2013. SAR image filtering based on the Cauchy–Rayleigh mixture model. *IEEE Geoscience and Remote Sensing Letters* 11, 960–964.
- Poulain, V., Inglada, J., Spigai, M., Tourneret, J.Y., Marthon, P., 2011. High-resolution optical and SAR image fusion for building database updating. *IEEE Transactions on Geoscience and Remote Sensing* 49, 2900–2910.
- R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Sagrillo, M., Guerra, R., Machado, R., Bayer, F., 2023. A generalized control chart for anomaly detection in SAR imagery. *Computers & Industrial Engineering* , 109030.
- Sagrillo, M., Guerra, R.R., Bayer, F.M., Machado, R., 2022. A new probability distribution for SAR image modeling. *Remote Sensing* 14, 2853.
- Soumekh, M., 1999. Synthetic aperture radar signal processing. volume 7. New York: Wiley.
- Vasconcelos, J.M., Cintra, R.J., Nascimento, A.D., Rego, L.C., 2020. The compound truncated Poisson Cauchy model: A descriptor for multimodal data. *Journal of Computational and Applied Mathematics* 378, 112887.
- Yue, D.X., Xu, F., Frery, A., Jin, Y.Q., 2020. A generalized Gaussian coherent scatterer model for correlated SAR texture. *IEEE Transactions on Geoscience and Remote Sensing* 58, 2947–2964.

---

## APÊNDICE A

# DESCRIÇÃO IMAGEM CARABAS-II

---



**Figura A.1:** Imagem SAR usada e mapa correspondente da mesma área.

**Tabela A.1:** Características da imagem SAR usada para a análise de dados.

Imagem Original	
Sistema	Carabas II
Origem	Agência Sueca de Pesquisa de Defesa
Tipo Imagem	Single-look
Missão	1
Banda	VHF
Polarização	Horizontal
Localização	Norte da Suécia 2002.
Dimensões	3000 × 2000 pixels

**Tabela A.2:** Características das regiões selecionadas.

	Região	
	Floresta	Lago
Dimensão	300 × 125 pixels	300 × 125 pixels
Posição	[101, 400] × [201, 325]	[2301, 2600] × [1851, 1975]