



Comparação de Modelagens Estatísticas para Dados SAR na Classificação de Cobertura da Terra: um Estudo de Caso na Região da Amazônia Brasileira

Comparison of Statistical Modeling for SAR Data in Land Cover Classification: a Case Study in the Brazilian Amazon Region

Vinicius D'Lucas Bezerra e Queiroz¹, Mariane Souza Reis², e Sidnei João Siqueira Sant'Anna³

¹Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brasil. vinicius.dlucas@gmail.com

ORCID: <https://orcid.org/0000-0002-9767-9982>

²Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brasil. mariane.reis@inpe.br

ORCID: <https://orcid.org/0000-0001-9356-7652>

³Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brasil. sidnei.santanna@inpe.br

ORCID: <https://orcid.org/0000-0001-7557-3215>

Recebido: 09.2024 | Aceito: 10.2025

Resumo: Imagens de Radar de Abertura Sintética (SAR) representam uma alternativa viável para a classificação da cobertura da terra em regiões frequentemente encobertas por nuvens. As amostras derivadas dessas imagens podem se adequar à diferentes distribuições estatísticas, dependendo do processamento realizado. Este trabalho objetivou comparar classificações de cobertura da terra obtidas com imagens SAR modeladas por diferentes distribuições estatísticas. Foram realizadas 1000 classificações supervisionadas de uma imagem PALSAR/ALOS em uma região do estado do Pará, considerando etapas típicas de pré-processamento, como filtragem, *multilooking* e correções geométricas. Utilizou-se o classificador de Máxima Verossimilhança adaptado às distribuições Gaussiana, Gama e Par de imagens em intensidade. As classes analisadas foram Agricultura, Solo Exposto, Pasto, Floresta e Vegetação Secundária. Os resultados indicaram que modelos com a mesma quantidade de polarizações apresentaram acurácias globais semelhantes, e que o uso de duas polarizações resultou em melhores acurácias do que a utilização de apenas uma. Neste trabalho, a distribuição Gaussiana, comumente adotada em classificadores paramétricos por simplificar a modelagem estatística, foi empregada na modelagem dos dados SAR sem comprometer o desempenho global da classificação, considerando as etapas de pré-processamento adotadas. Quanto às acurácias do produtor e do usuário, não foram encontrados modelos que separam bem todas as classes, com destaque para as classes Floresta e Vegetação Secundária, cuja separabilidade é difícil. Entretanto, a classe Solo Exposto apresentou boa distinção em todos os modelos.

Palavras-chave: Classificação. Cobertura da terra. SAR. Modelagens estatísticas.

Abstract: Synthetic Aperture Radar (SAR) images are a viable alternative for land cover classification in areas frequently covered by clouds. Samples from these images may fit to different statistical distributions depending on the preprocessing applied. This study aimed to compare land cover classifications obtained from SAR images modeled with different statistical distributions. A total of 1000 supervised classifications were performed on a PALSAR/ALOS image from a region in Pará State, Brazil, considering typical SAR preprocessing steps such as filtering, multilooking, and geometric corrections. The Maximum Likelihood classifier was adapted to Gaussian, Gamma, and the Joint Distribution of intensity images. The analyzed classes were Agriculture, Bare Soil, Pasture, Forest, and Secondary Vegetation. Results indicated that models using the same number of polarizations had similar overall accuracies, and that using two polarizations resulted in better accuracies than using a single polarization. In this study, the Gaussian distribution, commonly adopted in parametric classifiers because it simplifies statistical modeling, was used in SAR data modeling without compromising the overall classification performance, considering the image preprocessing steps adopted. Regarding producer's and user's accuracies, none of the models achieved a clear separation among all classes, especially the Forest and Secondary Vegetation classes, which are difficult to separate. However, the Exposed Soil class showed good distinction in all models.

Keywords: Classification. Land cover. SAR. Statistical modeling.

1 INTRODUÇÃO

A capacidade de aquisição de imagens de grandes áreas e em intervalos regulares de tempo faz do Sensoriamento Remoto (SR) uma importante ferramenta para obtenção de informações contínuas sobre o meio ambiente (Jensen, 2014; Schowengerdt, 2007). Por meio de imagens, é possível extrair informações sobre a cobertura da terra e fazer inferências sobre o seu uso. Neste contexto, a cobertura se refere ao caráter físico e biótico da superfície terrestre, enquanto o uso está relacionado à finalidade para qual a cobertura é utilizada pelo ser humano (Lambin et al., 2006; Turner & Meyer, 1994).

Imagens de sensores ópticos são comumente utilizadas para classificar a cobertura da terra (Martins et al., 2018; Wulder et al., 2015). Entretanto, um desafio relacionado a essa tarefa é a presença de nuvens na cena, as quais podem inviabilizar a classificação dos alvos diretamente em classes de cobertura da terra. Uma alternativa ao problema da presença de nuvens na cena é a utilização de imagens de sensores de Radar de Abertura Sintética (*Synthetic Aperture Radar*, SAR). Esses sensores operam na região espectral das micro-ondas (1 mm a 1 m) e permitem, na maioria dos casos, o imageamento através das nuvens em virtude da alta transmitância atmosférica nessa faixa espectral (Woodhouse, 2006). Essa propriedade torna os dados SAR especialmente úteis em condições meteorológicas adversas em que há cobertura persistente de nuvens (Meyer, 2019; Richards, 2009). Além disso, os sensores SAR dispõem de fonte ativa e controlável de radiação, o que possibilita o imageamento independente da iluminação solar.

Os mapas de uso e cobertura podem ser gerados por meio da classificação de imagens. Um dos métodos paramétricos tradicionais de classificação supervisionada mais utilizados com imagens de SR é o classificador de Máxima Verossimilhança (Maxver) (Richards, 2022; Schowengerdt, 2007). As principais implementações desse método adotam frequentemente a suposição de que as amostras das classes possuem uma distribuição Gaussiana – o que simplifica a modelagem estatística e a inferência, além de permitir a obtenção de medidas probabilísticas de pertencimento de cada *pixel* às classes de interesse. Quando aplicado a dados SAR, entretanto, surge uma limitação importante: a adequação da distribuição estatística utilizada. Amostras das classes de cobertura adquiridas em dados SAR frequentemente não seguem uma distribuição Gaussiana. Nesse sentido, diversos estudos na literatura demonstraram que os dados SAR podem ser mais bem representados por diferentes distribuições estatísticas, a depender do formato do dado, nível de processamento, dimensionalidade e grau de homogeneidade dos alvos na cena (Frery, Correa et al., 1999; Frery et al., 1997, 2007; Gao, 2010; Lee & Pottier, 2017; Mahapatra et al., 2024; Oliver & Quegan, 2004; Yue et al., 2021).

No entanto, etapas de pré-processamento realizadas na imagem SAR podem alterar significativamente suas propriedades estatísticas – sobretudo aquelas voltadas à redução do ruído *speckle*, como o *multilooking* e a filtragem. Nesses casos, pelo efeito do Teorema Central do Limite (TCL) (Kwak & Kim, 2017), os dados podem se aproximar do comportamento de variáveis aleatórias normalmente distribuídas. A investigação comparativa ou a proposição de novos modelos estatísticos permanecem como linhas de pesquisa em evolução, no que se refere à análise de aderência dos dados SAR a distintos modelos (Negri et al., 2016a; Sagrillo et al., 2022; Yue et al., 2021).

Como diferentes distribuições estatísticas podem ser adequadas ao processamento de dados SAR, é comum que pesquisadores menos experientes com esse tipo de dado suponham, de forma equivocada, a necessidade de adotar uma distribuição específica. Em alguns casos, essa distribuição nem sempre está implementada em *softwares* mais acessíveis. Além disso, essa percepção pode levar ao uso de classificadores não paramétricos, que exigem conjuntos de amostras distintos, especialmente quando a amostragem foi ajustada para abordagens paramétricas. Desta forma, o objetivo deste trabalho é avaliar o impacto do uso de diferentes distribuições estatísticas quando utilizadas na classificação de imagens SAR.

Neste estudo, optou-se por utilizar o classificador de Máxima Verossimilhança (Maxver). Esse classificador foi escolhido por sua robustez e facilidade de implementação e aplicação. Além de ser um classificador comumente presente na maioria dos *softwares* de processamento de imagens de SR, o Maxver também permite incorporar funções de densidade de probabilidade, tornando-o particularmente apto para o objetivo do estudo. As distribuições estatísticas Gama (caso univariado, i.e., uma polarização) e Par de imagens intensidade (caso bivariado, i.e., duas polarizações) foram incorporadas ao classificador Maxver para avaliar se produzem melhores

resultados de classificação do que a distribuição Gaussiana. Foi considerado, em especial, o emprego de etapas de pré-processamento comumente utilizadas em aplicações com imagens SAR, como *multilooking*, filtragem de *speckle* e correção geométrica. A questão central investigada é se, diante desse cenário, ainda se justifica a utilização de distribuições consideradas próprias de imagens SAR ou se a hipótese Gaussiana permanece suficiente para a modelagem estatística. Como objetivo secundário, esse trabalho também fornece um conjunto de ferramentas para a análise de dados SAR, em linguagem R.

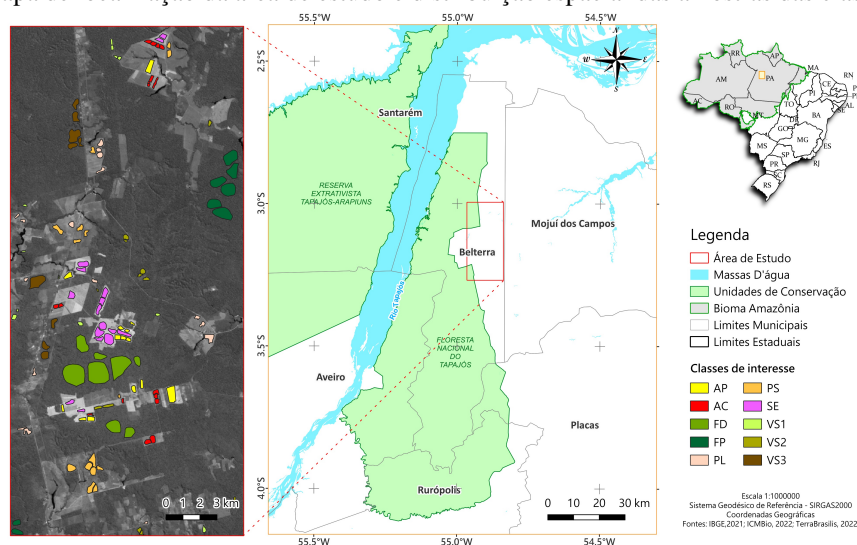
Para o teste da metodologia proposta, procedeu-se à classificação de uma imagem do sensor *Phase Array L-Band Synthetic Aperture Radar* (PALSAR), a bordo do satélite *Advanced Land Observing System* (ALOS), de uma região do município de Belterra, localizada no estado do Pará, na Amazônia brasileira. Métricas de avaliação foram conduzidas em nível global e de classe, de modo a examinar também as diferenças de desempenho entre distintas classes de cobertura.

2 MATERIAL

2.1 Área de estudo e classes de interesse

A área de estudo está localizada no município de Belterra, no Pará (Figura 1) e intersecta parte da Floresta Nacional (FLONA) do Tapajós, uma Unidade de Conservação Federal criada pelo Decreto Federal n.º 73.684/1974 (Brasil, 1974). Essa região apresenta clima característico de florestas tropicais úmidas, com temperatura média anual de 26 °C e precipitação pluviométrica anual entre 1700 e 2000 mm (Servello, 2010). A vegetação original é do tipo Floresta Ombrófila Densa, com abundância e ampla distribuição de indivíduos arbóreos de grande porte, lianas lenhosas, palmeiras e epífitas (Veloso et al., 1991).

Figura 1 – Mapa de localização da área de estudo e distribuição espacial das amostras das classes de interesse.



Fonte: Os autores (2025).

As transformações ocasionadas pelo processo de ocupação intensivo nessa região proporcionaram uma variedade de classes de uso e cobertura da terra, despertando interesse como área de estudo para condução de pesquisas científicas nas últimas décadas (Negri et al., 2016; Pereira et al., 2018; Reis et al., 2018, 2024, 2020a; C. S. F. Yanasse et al., 1997). Assim, a complexidade da paisagem motivou a seleção dessa região como área de estudo, bem como: a) a disponibilidade de imagens SAR; b) a disponibilidade de dados de referência de (Sant'Anna et al., 2016); c) conhecimento prévio de classes de uso e cobertura da terra; e d) a existência de legenda padronizada de classes de uso e cobertura da terra (Reis et al., 2018, 2024).

As classes de uso e cobertura utilizadas na classificação foram definidas, preliminarmente, com base em um modelo conceitual e hierárquico de classes desenvolvido para a área de estudo (Reis et al., 2018). Foram identificadas 10 classes de cobertura em imagens ópticas para a área de estudo no ano de 2010, descritas no Quadro 1. A disposição espacial das amostras coletadas para essas classes pode ser observada na Figura 1 sobre

uma imagem óptica datada de 28 de junho de 2010, da banda do infravermelho próximo (Banda 4) do sensor *Thematic Mapper* (TM) a bordo do satélite Landsat-5.

Quadro 1 – Descrição das classes de cobertura da terra.

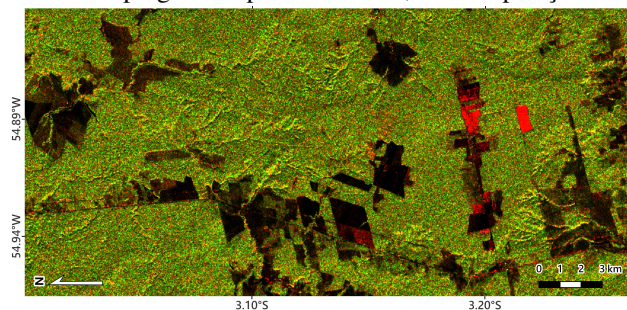
Classe	Descrição
Área Cultivada (AC)	Culturas de grãos.
Áreas Agrícolas em Pousio (AP)	Áreas agrícolas em pousio cobertas por palha ou vegetação esparsa.
Floresta Primária (FP)	Floresta na qual a ação humana não provocou alterações significativas em suas características originais de estrutura e espécies.
Floresta Modificada (FD)	Floresta degradada por atividades de fogo ou por desmatamento seletivo cujas características originais tenham sido alteradas, mas ainda possuam porte e fisionomia florestal.
Pasto Limpo (PL)	Áreas com vegetação típica de pastagens com predomínio de herbáceas.
Pasto Sujo (PS)	Áreas com vegetação típica de pastagens com presença de espécies arbustivas e espécies invasoras.
Solo Exposto Agrícola (SE)	Áreas predominantemente de solo exposto com destinação agrícola.
Vegetação Secundária Inicial (VS1)	Áreas de vegetação secundária com predominância de herbáceas e arbustos.
Vegetação Secundária Intermediária (VS2)	Áreas com presença de vegetação secundária com predominância de árvores de pequeno porte e arbustos.
Vegetação Secundária Avançada (VS3)	Florestas secundárias em avançado estágio de desenvolvimento com predomínio de árvores e ocorrência de árvores emergentes, arbustos e herbáceas.

Fonte: Adaptado de Reis et al. (2018).

2.2 Imagens e recursos computacionais utilizados

Este estudo se baseia em uma imagem adquirida em 21 de junho de 2010 (Figura 2) pelo sensor *Phase Array L-Band Synthetic Aperture Radar* (PALSAR), a bordo do satélite *Advanced Land Observing System* (ALOS), desenvolvido pela *Japan Aerospace Exploration Agency* (JAXA). A imagem foi obtida por meio do repositório *Alaska Satellite Facilities* (ASF) *data search*, no modo de aquisição *Fine Beam Dual* (FBD), com nível de processamento 1.1 e fornecida no formato *Single Look Complex* (SLC). O sensor opera na banda L (23,6 cm) e, para este modo de aquisição, adquire informações nas polarizações HH e HV. A resolução espacial é de aproximadamente 10 m na direção do alcance (*range*) e 4,5 m na direção do azimute, com espaçamento entre *pixels* de cerca de 9,36 m e 3,19 m, respectivamente (Wright et al., 2009).

Figura 2 – Imagem PALSAR/ALOS empregada no processamento, em composição colorida HH (R), HV (G) e HH (B).



Fonte: Os autores (2025).

Os recursos computacionais empregados neste estudo compreenderam diferentes ambientes e ferramentas especializadas. O QGIS (versão 3.22.13) foi utilizado para a visualização, manipulação de dados georreferenciados e elaboração de figuras. O *software* SNAP/STEP (versão 9.0.0) foi empregado nas etapas de pré-processamento das imagens. Para as rotinas de processamento de imagens, cálculos estatísticos e geração de gráficos, utilizou-se o ambiente *RStudio* (versão 2023.06.0+421, *Mountain Hydrangea*) em conjunto com o R (versão 2.4.5).

3 MÉTODOS

A metodologia adotada neste trabalho está dividida nas seguintes etapas: pré-processamento, modelagem estatística, seleção de amostras, cálculo do número equivalente de *looks* (NEL), análise de separabilidade das classes, classificação supervisionada e avaliação dos resultados. Cada uma dessas etapas é detalhada nas seções subsequentes.

3.1 Pré-processamento

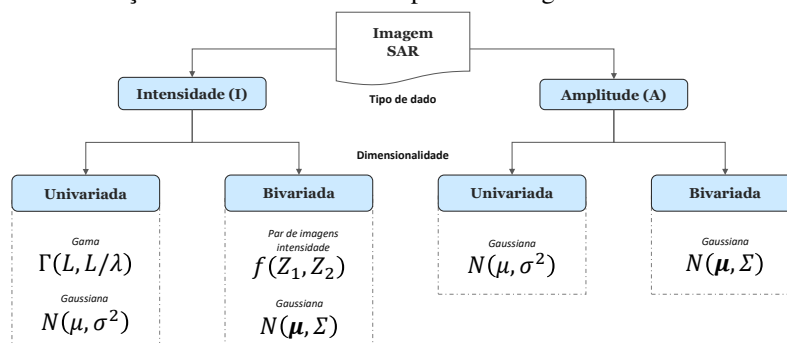
O pré-processamento da imagem PALSAR/ALOS incluiu as etapas de: calibração, *multilooking*, filtragem *speckle*, *deskewing* e correção do terreno, com base no tutorial de ortorretificação de imagens PALSAR/ALOS da Agência Espacial Europeia (ESA) (Veci, 2015). As imagens PALSAR/ALOS em nível de processamento 1.1 foram, primeiramente, calibradas radiometricamente em coeficiente de retroespalhamento (σ^0) em escala linear (European Space Agency, 2007; Shimada et al., 2009). Em seguida, as imagens passaram pelo processo de *multilooking*, visando gerar *pixels* de dimensões aproximadamente iguais e reduzir o efeito do *speckle*. Para isso, realizou-se o promédio de valores de *pixels* adjacentes na direção de azimuth, equivalendo a cerca de cinco *looks* nessa direção. Após esse processo, o filtro *Gamma-MAP* foi aplicado com uma janela móvel de 3×3 *pixels*.

Por fim, foi realizado o processo de correção geométrica das imagens, que inclui o *deskewing* e a correção do terreno. O *deskewing* é um processo recomendado no processamento de imagens PALSAR/ALOS, que transforma o dado para uma geometria zero-Doppler e tem o objetivo de corrigir distorções sistemáticas na geometria da imagem (Garg & Singh, 2018). A correção do terreno foi efetuada pelo método *range-Doppler* utilizando como modelo digital de elevação o *Shuttle Radar Topography Mission* (SRTM) de resolução de 3 segundos de arco. Como resultado desse fluxo de procedimentos foram obtidas imagens em intensidade e amplitude nas polarizações HH e HV calibradas radiometricamente e geocodificadas, com espaçamento de *pixel* de 15 m.

3.2 Modelagens estatísticas

Diferentes distribuições estatísticas foram empregadas para a modelagem das imagens SAR de acordo com a dimensionalidade e o tipo dos dados (Figura 3). No que se refere à dimensionalidade, os dados univariados consistem nas duas polarizações da imagem PALSAR/ALOS (HH e HV) modeladas e classificadas individualmente. Em contrapartida, nos casos bivariados, as duas polarizações são utilizadas simultaneamente como variáveis na modelagem. As amostras das classes foram modeladas com as distribuições Gama e Gaussiana no caso univariado, e com as distribuições de Par de imagens intensidade e Gaussiana multivariada, no caso bivariado. Para os dados em amplitude, as classificações foram realizadas com modelagens Gaussianas tanto no caso univariado quanto no bivariado.

Figura 3 – Distribuições estatísticas utilizadas para modelagens de amostras da imagem SAR.



Fonte: Os autores (2025).

3.2.1 DISTRIBUIÇÃO GAMA

Dentre as diversas distribuições estatísticas que podem caracterizar os dados SAR, para imagens SAR *multilook*, com uma polarização e em formato intensidade, a distribuição Gama modela a intensidade de retorno $Z_I \sim \Gamma(L, L/\lambda)$ de regiões homogêneas (Frery, Freitas et al., 1999; Frery et al., 1997; Gao, 2010). Nessas regiões, o retroespalhamento X é considerado constante ($X \sim \lambda > 0$). Assim, a intensidade do sinal de retorno Z_I tem sua função de densidade dada por:

$$f_{Z_I}(z; L, \lambda) = \frac{L^L}{\lambda^L \Gamma(L)} z^{L-1} e^{-\frac{Lz}{\lambda}}, \quad \text{para } L \geq 1 \text{ e } z, \lambda > 0, \quad (1)$$

em que L representa o número de *looks* da imagem SAR e λ uma função constante. Assim, a Equação 1 pode ser utilizada como função de densidade de probabilidade condicional para cada classe de cobertura em classificadores probabilísticos em substituição à abordagem Gaussiana.

3.2.2 DISTRIBUIÇÃO PAR DE IMAGENS INTENSIDADE

A distribuição para um par de imagens intensidade é derivada da distribuição *Wishart* complexa multivariada (Lee et al., 1994), apresentando grande importância na modelagem de dados SAR polarimétricos incompletos. Considerando duas imagens SAR *multilook* em intensidade, em que seus sinais de retorno são representados por Z_1 e Z_2 , a função de densidade de probabilidade conjunta entre essas duas imagens é dada por (Lee et al., 1994):

$$f_{Z_1, Z_2}(z_1, z_2, h_{11}, h_{22}, L, \rho_c) = \frac{L^{L+1} (z_1 z_2)^{\frac{L-1}{2}} \exp \left[-\frac{L(z_1/h_{11} + z_2/h_{22})}{1 - |\rho_c|^2} \right]}{(h_{11} h_{22})^{\frac{L+1}{2}} \Gamma(L) (1 - |\rho_c|^2) |\rho_c|^{L-1}} I_{L-1} \left(\frac{2L|\rho_c|}{1 - |\rho_c|^2} \sqrt{\frac{z_1 z_2}{h_{11} h_{22}}} \right), \quad (2)$$

em que

$$h_{11} = \mathbb{E}[Z_1], \quad (3)$$

$$h_{22} = \mathbb{E}[Z_2], \quad (4)$$

e $|\rho_c|^2$ é dado pela equação

$$|\rho_c|^2 = \frac{\mathbb{E}[(Z_1 - h_{11})(Z_2 - h_{22})]}{\sqrt{\mathbb{E}[(Z_1 - h_{11})^2] \mathbb{E}[(Z_2 - h_{22})^2]}}. \quad (5)$$

Os termos $|\rho_c|$ e θ representam o módulo e a fase do coeficiente de correlação complexo, respectivamente, e I_{L-1} é a função de *Bessel* modificada de ordem $L - 1$. As estimativas dos parâmetros h_{11} , h_{22} e $|\rho_c|$ podem ser encontradas em Correia (1999).

Assim como a distribuição Gama pode ser utilizada para o caso univariado, é possível utilizar a distribuição para um par de imagens de intensidade como função de densidade de probabilidade condicional na modelagem das amostras das classes para o caso bivariado. É importante ressaltar que essa condição é válida supondo que as amostras coletadas são de regiões que possuem um comportamento homogêneo em relação ao seu sinal de retroespalhamento, o que caracteriza a condição de *fully-developed speckle* (Yue et al., 2021), e não simplesmente amostras pertencentes à mesma classe.

3.3 Seleção de amostras

Foram utilizadas amostras produzidas por Reis et al. (2017) (Figura 1) para a classificação das imagens SAR referentes ao ano de 2010. Essas amostras foram coletadas por meio de interpretação visual de imagens ópticas adquiridas em datas próximas às das imagens SAR, e possuem validação com informações adquiridas em campo e confrontação com dados do projeto de Monitoramento Sistemático da Cobertura e Uso da Terra (TerraClass) (Almeida et al., 2016). Considerando que os sensores SAR registram propriedades distintas dos alvos em comparação aos sensores ópticos, a interpretação visual foi realizada em imagens ópticas de modo a garantir maior interpretabilidade e consistência no procedimento de seleção de amostras. Assim, a partir desse conjunto amostral, foi realizada uma análise de autocorrelação espacial nas direções horizontal e vertical entre os *pixels* das amostras das 10 classes previamente definidas, calculando os *lags* (ou deslocamento em número de *pixels*) mínimos em que as amostras estejam descorrelacionadas espacialmente.

Dada uma amostra retangular de uma imagem, representada como $z = [z_{i,j}]_{1 \leq i \leq N_x, 1 \leq j \leq N_y}$, em que N_x e N_y indicam, respectivamente, o número de amostras (*pixels*) ao longo das direções horizontal (x) e vertical (y) do retângulo e $N = N_x \times N_y$ é o total de amostras contidas nesse retângulo, define-se $a = [z_{i,j} - \bar{z}]_{1 \leq i \leq N_x, 1 \leq j \leq N_y}$

como o conjunto de valores centralizados em torno da média \bar{z} , de modo que $\bar{a} = 0$. Sendo s_x e s_y ($s_x, s_y \in \mathbb{N}$) os lags nas direções x e y , a estimativa da autocovariância espacial é dada por (C. C. F. Yanasse, 1991):

$$\hat{\gamma}_a(s_x, s_y) = \frac{1}{N} \sum_{t_x=k_x}^{K_x} \sum_{t_y=k_y}^{K_y} a_{t_x, t_y} a_{t_x+s_x, t_y+s_y}, \quad (6)$$

em que $k_x = \max\{1, 1 - s_x\}$, $k_y = \max\{1, 1 - s_y\}$, $K_x = \min\{N_x, N_x - s_x\}$ e $K_y = \min\{N_y, N_y - s_y\}$. A estimativa da autocorrelação espacial é definida como $\hat{\rho}_a = \hat{\gamma}_a / \hat{\gamma}_a(0, 0)$.

O lag adotado para seleção de amostras consistiu no menor número natural no qual a autocorrelação espacial é menor ou igual a 0,3. Esse procedimento foi realizado para cada polígono de amostras nas duas polarizações da imagem PALSAR/ALOS. Para gerar um conjunto amostral único, foi selecionado o maior valor de lag entre as polarizações para cada classe, desconsiderando outliers. Como resultado, foram obtidos grids de amostras por classes regularmente espaçadas pelos valores de lags e, portanto, descorrelacionadas. Após descorrelação, as amostras foram utilizadas para treinamento e validação dos modelos de classificação neste trabalho.

3.4 Cálculo do Número Equivalente de Looks (NEL)

Após a seleção amostral, uma metodologia foi definida para o cálculo do Número Equivalente de looks (NEL) da imagem SAR. O NEL representa a estimativa do número de looks da imagem SAR e é utilizado para descrever a qualidade do dado original no seu processo de formação, assim como para caracterizar a suavização e a redução do speckle em operações de processamento de imagens, tais como a filtragem (Anfinset et al., 2009). O valor do NEL está inversamente relacionado com o grau de speckle existente na imagem, ou seja, quanto maior o NEL, menor o efeito desse fenômeno (Lee & Pottier, 2017).

Sabendo que a distribuição Gama modela o sinal de retorno de regiões homogêneas da imagem SAR-multilook em intensidade e que a modelagem de todas as classes utiliza o mesmo L como um dos parâmetros de sua distribuição, admitiu-se que o L estimado fosse aquele cuja classe: a) tenha suas amostras aceitas no teste de aderência de Kolmogorov-Smirnov (Kolmogorov, 1933; Smirnov, 1939) com 5% de nível de significância; b) produza o maior número de classes aceitas no teste quando o seu parâmetro de forma é utilizado para remodelar as outras classes; e c) no caso de empates no item anterior, observa-se o maior p -valor dentre as classes aceitas no teste.

Os NELs foram calculados para as polarizações HH e HV da imagem PALSAR/ALOS em intensidade. Como a imagem em amplitude foi modelada somente com distribuição Gaussiana, não foi necessário estimar o NEL para esse dado.

3.5 Análise da separabilidade das classes

Análise da separabilidade das classes foi realizada com o objetivo de verificar quais delas poderiam ser agregadas, sob os pontos de vista estatístico e conceitual, de modo a reduzir a confusão entre classes, aumentar a discriminabilidade e, conseqüentemente, melhorar os resultados de classificação. Assim, a métrica de separabilidade calculada foi a distância estocástica de Bhattacharyya modelada para as distribuições Gaussiana e Gama. Em termos práticos, quanto maior a distância calculada, menor a sobreposição e mais clara é a distinção entre as duas distribuições. As expressões para o cálculo das distâncias de Bhattacharyya para as distribuições Gaussiana multivariada $d_{B(N)}$ (Theodoridis & Koutroumbas, 2009) e Gama $d_{B(\Gamma)}$ (Torres, 2012) são:

$$d_{B(N)}(\mu_1, \Sigma_1, \mu_2, \Sigma_2) = \frac{1}{8} \left[(\mu_1 - \mu_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (\mu_1 - \mu_2) \right] + \frac{1}{2} \ln \left(\frac{\left| \frac{\Sigma_1 + \Sigma_2}{2} \right|}{\sqrt{|\Sigma_1| |\Sigma_2|}} \right) \quad (7)$$

$$d_{B(\Gamma)}(\mu_1, \mu_2) = \ln \left(\frac{(\mu_1 + \mu_2)^L}{2^L (\mu_1 \mu_2)^{L/2}} \right), \quad (8)$$

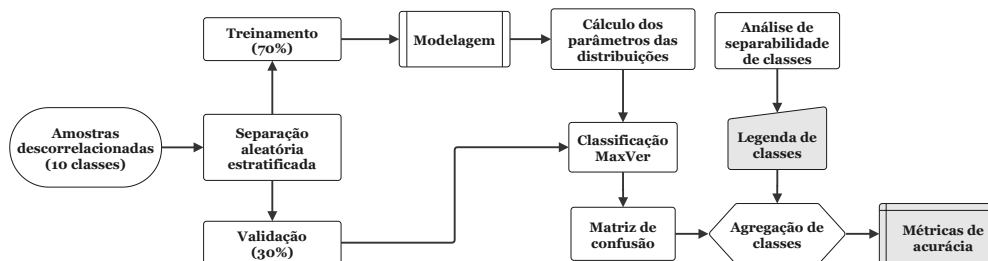
em que μ representa a média, Σ a matriz de covariância, L indica o NEL da imagem, $(\cdot)^{-1}$ simboliza a inversa da matriz, $|\cdot|$ indica o determinante da matriz e $(\cdot)^\top$, a transposta da matriz. A distância $d_{B(\Gamma)}$ é a distância de *Bhattacharyya* entre duas distribuições Gama com o mesmo parâmetro de forma L .

Para todos os pares de classes, foram calculadas as distâncias *Bhattacharyya*. No caso dos dados SAR univariados em intensidade, foi utilizada a distância de *Bhattacharyya* entre distribuições Gama, enquanto para os dados SAR em amplitude e para o caso bivariado, as distâncias foram modeladas com distribuição Gaussiana. A partir disso, foram criados dendrogramas empregando o método de *Ward* (Ward Jr., 1963) para auxiliar na visualização da separabilidade das classes e da estrutura dos agrupamentos hierárquicos formados. Esses dendrogramas, juntamente com a legenda conceitual de classes descrita na Seção 2.1, foram utilizados para desenvolver a legenda de classes aplicada na classificação de cobertura.

3.6 Estratégias de classificação e avaliação dos resultados

Após definir a amostragem, os parâmetros das distribuições e a legenda de classes, o processo de classificação da imagem SAR foi iniciado (Figura 4). As amostras de treinamento e validação provenientes do conjunto descorrelacionado foram selecionadas de forma aleatória e estratificada, com 70% destinadas ao treinamento e os 30% restantes à avaliação das classificações. Um total de 1000 classificações foram realizadas para cada modelo utilizando o classificador de Máxima Verossimilhança (Maxver) adaptado às distribuições em estudo, variando as amostras dos conjuntos de treinamento e validação nas proporções mencionadas, com reposição.

Figura 4 – Fluxograma da estratégia de classificação empregada no estudo. Os itens em cinza simbolizam produtos resultantes do fluxo.



Fonte: Os autores (2025).

A cada iteração, a classificação foi inicialmente realizada com 10 classes de cobertura de nível hierárquico mais baixo. Posteriormente, essas classes foram agregadas na matriz de confusão para a legenda final (resultado da Seção 3.5). A escolha de uma estratégia de agregação pós-classificação visou evitar interferências na modelagem e na estimativa inicial dos parâmetros das classes. Para cada matriz de confusão gerada, foram calculadas as métricas de Acurácia Global (A.G.), Acurácia do Usuário (A.U.) e Acurácia do Produtor (A.P.). Por fim, foram calculadas as médias e o intervalo de credibilidade (IC) de 95% para cada métrica, excluindo os 2,5% dos maiores e menores valores encontrados. Como resultados, foram gerados gráficos comparativos das métricas globais e por classe de cada modelo, bem como a moda dos 1000 resultados de classificação.

4 RESULTADOS

4.1 Análise exploratória das amostras

A Tabela 1 apresenta os *lags* calculados nas direções vertical e horizontal, bem como a quantidade e proporção de amostras por classe depois da análise de autocorrelação espacial. A partir desta tabela, nota-se

um desbalanceamento no número de amostras (*pixels*) por classe. Esse desbalanceamento está associado à disponibilidade das classes na cena, bem como ao método de coleta empregado.

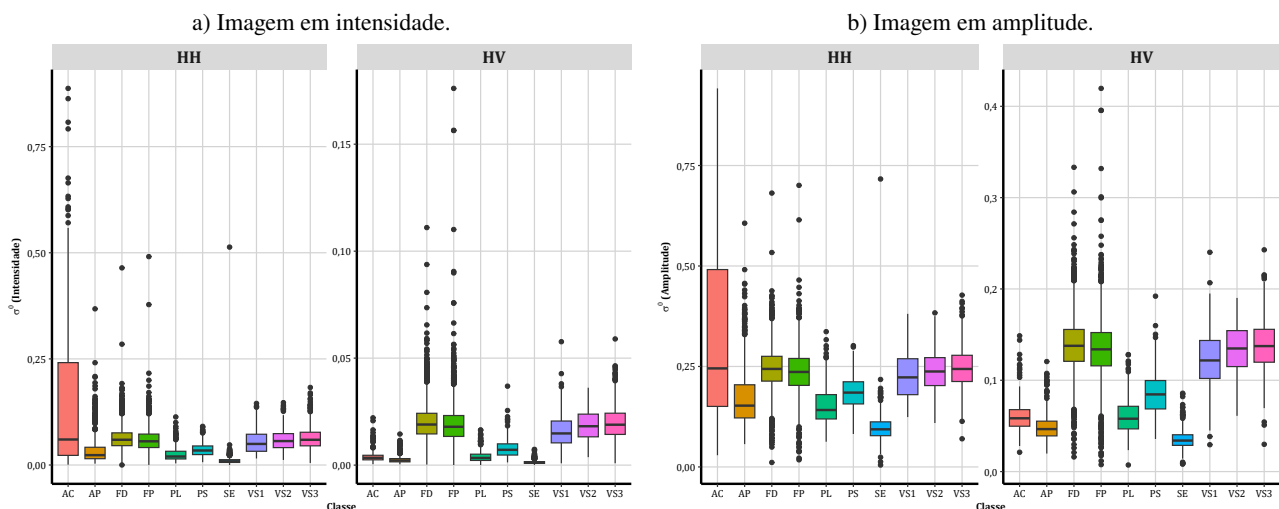
Tabela 1 – Valores de *lags* nas direções horizontal (H) e vertical (V) calculados para as classes de interesse. Os símbolos # e % representam, respectivamente, o número e a proporção de amostras por classe.

	AP	AC	FD	FP	PL	PS	SE	VS1	VS2	VS3	Total
<i>lag H</i>	3	4	2	2	4	4	3	4	3	3	-
<i>lag V</i>	4	4	3	2	3	4	3	4	4	3	-
#	259	535	4156	3449	413	454	1029	121	216	927	11559
%	2,24	4,63	35,95	29,84	3,57	3,93	8,90	1,05	1,87	8,02	100

Fonte: Os autores (2025).

Os valores das amostras para os dados em intensidade e em amplitude são apresentados na Figura 5. Nota-se que valores mais altos do sinal de retorno predominam na polarização HH. Além disso, classes associadas a coberturas florestais, como Floresta Primária (FP), Floresta Modificada (FD), e Vegetações Secundárias Inicial (SV1), Intermediária (SV2) e Avançada (SV3), exibem uma distribuição semelhante e medianas com valores próximos, o que dificulta a separação na classificação de imagens. Visivelmente, a classe de Área Cultivada (AC) se distingue das demais pela variância dos seus dados na polarização HH, sugerindo um comportamento distinto das demais classes nessa polarização. Em termos de separabilidade das classes, a polarização HV mostra uma melhor discriminação para os dois tipos de dados, evidenciada por menores sobreposições dos intervalos interquartis de cada *boxplot* se comparados aos da polarização HH.

Figura 5 – Distribuição do sinal de retorno por classe e polarização, com escalas ajustadas para visualização e cores de classe atribuídas aleatoriamente.



Fonte: Os autores (2025).

4.2 Cálculo do Número Equivalente de Looks (NEL)

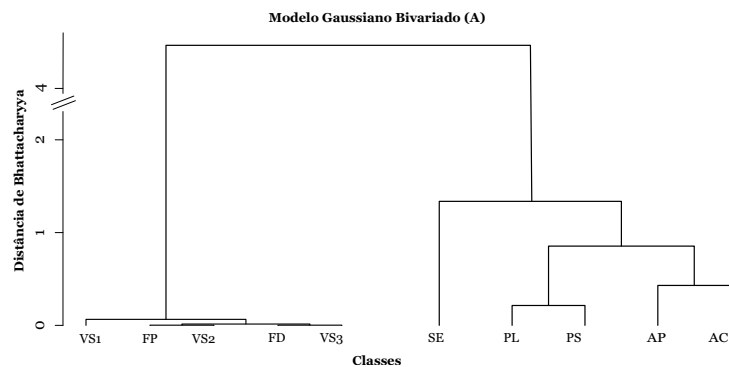
Seguindo a metodologia proposta para o cálculo dos NELs para as polarizações HH e HV, os parâmetros das classes foram estimados utilizando a distribuição Gama. Dentre esses parâmetros, aqueles que resultaram no maior número de classes aceitas no teste de aderência foram os correspondentes às classes Pasto Sujo (PS) ($\hat{L} = 5,591$) para a polarização HH e Solo Exposto Agrícola (SE) ($\hat{L} = 3,690$) para a polarização HV. Esses valores foram, portanto, adotados como NEL para as respectivas polarizações. Para distribuição Par de imagens intensidade, a média desses valores foi utilizada como NEL ($\hat{L} = 4,640$).

4.3 Análise de separabilidade das classes e definição de legenda final

Os resultados das distâncias *Bhattacharyya* entre os *clusters* de amostras das classes foram representados em dendrogramas. A Figura 6 exemplifica um dos dendrogramas produzidos para avaliação da separabilidade das classes. Nela, é possível observar a baixa dissimilaridade entre as classes FP, FD, VS1, VS2 e VS3. Especialmente,

no modelo Gaussiano bivariado em amplitude, constata-se a coerência na ordem de agregação das classes, sendo a classe SE a que é melhor discriminada.

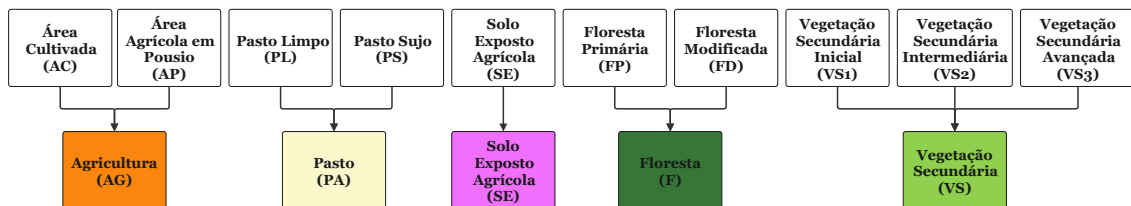
Figura 6 – Exemplo de dendrograma utilizado para avaliação da separabilidade das classes.



Fonte: Os autores (2025).

Analisando os padrões de agregação apresentados nos dendrogramas, juntamente com a análise exploratória das amostras, foi proposta uma nova legenda com cinco classes: Agricultura (AG), Pasto (PA), Floresta (F), Solo Exposto Agrícola (SE) e Vegetação Secundária (VS). Optou-se por manter separadas classes com alta separabilidade e também as que possuem baixa separabilidade mas que representam alvos no solo com significados muito distintos (e.g, Floresta Primária vs. Secundária). A nova configuração das classes, que resultou na legenda utilizada para as classificações da imagem SAR, é apresentada na Figura 7.

Figura 7 – Legenda final de classes de cobertura.



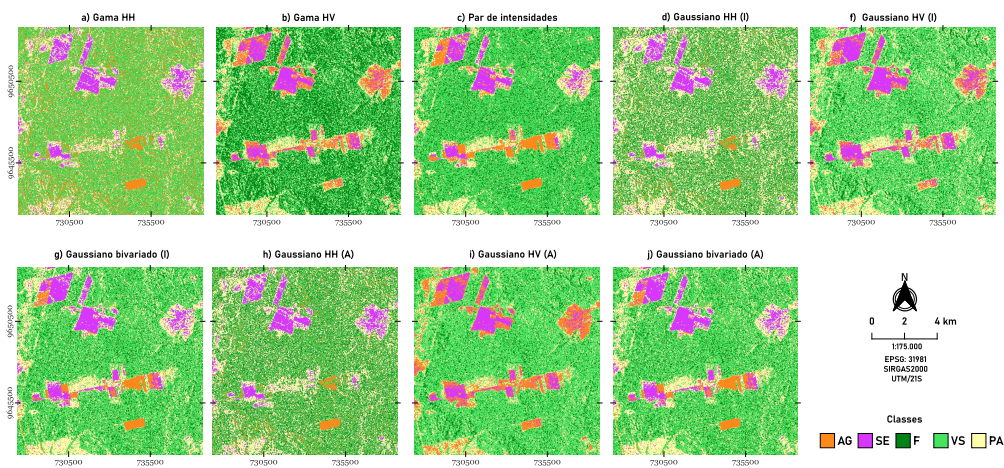
Fonte: Os autores (2025).

4.4 Avaliação de desempenho dos modelos de classificação

A moda de cada conjunto de 1000 classificações pode ser observada na Figura 8, enquanto os valores das métricas que medem a exatidão dos resultados, tanto no nível global quanto por classe, estão apresentados na Figura 9. Os modelos bivariados apresentaram as melhores médias de Acurácia Global (A.G.), em que se destaca o valor para distribuição Par de imagens intensidade (A.G. = 0,59). As piores médias estão relacionadas com os modelos que utilizaram apenas a polarização HH, destacando-se o modelo Gama HH, cuja A.G. média foi de 0,45. Os modelos que utilizaram apenas a polarização HV apresentaram melhores acurácias globais médias que aqueles que utilizaram apenas a polarização HH. As sobreposições dos ICs indicam desempenhos similares das modelagens e ocorrem geralmente em modelos com mesma dimensionalidade e polarizações utilizadas.

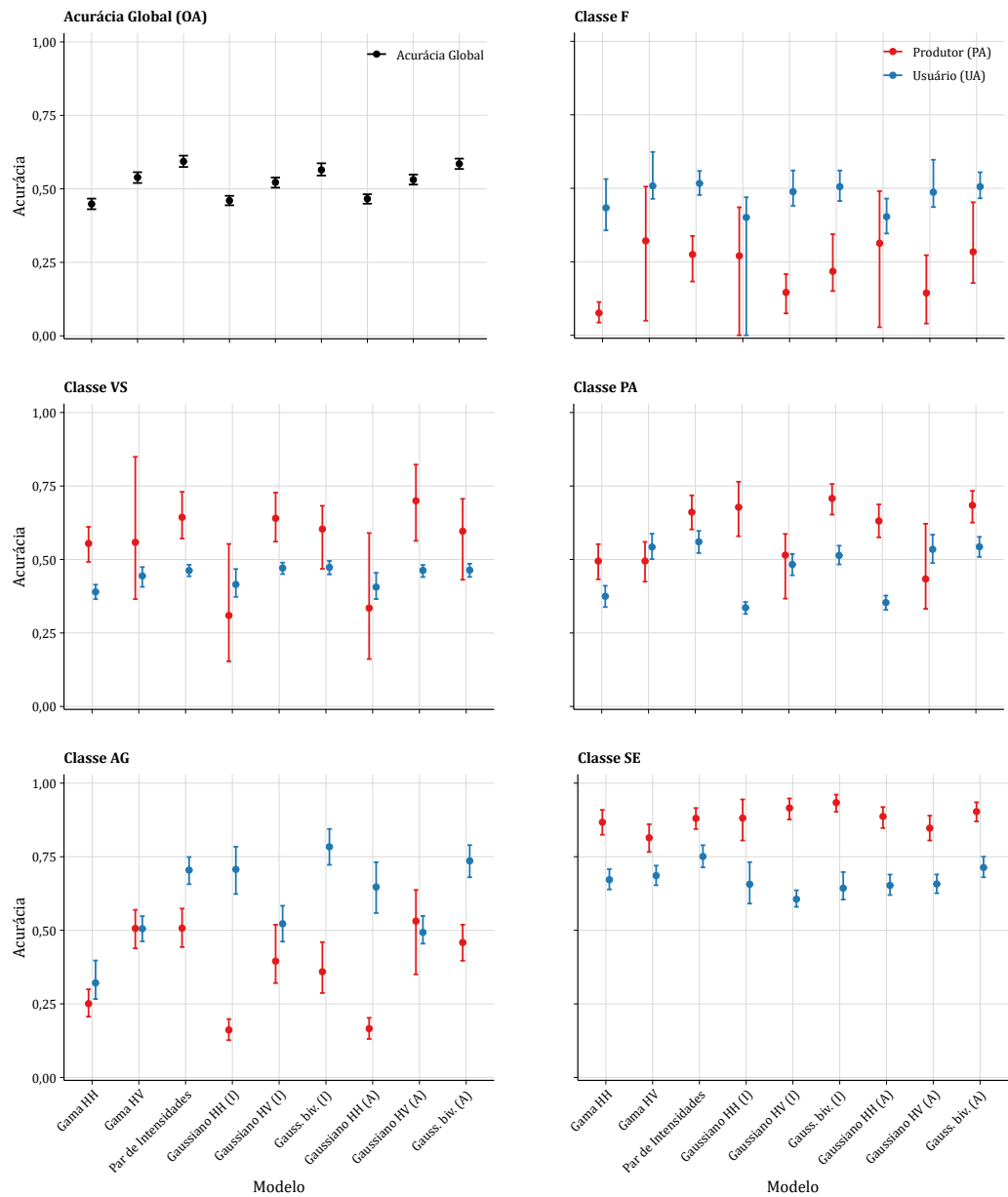
Em relação ao comportamento por classe, AG apresentou piores valores de Acurácia do Produtor (A.P.) nos modelos com apenas polarização HH, enquanto os demais modelos exibiram semelhanças estatísticas. Para PA, o modelo Gaussiano bivariado em intensidade alcançou o maior valor médio (0,71), embora sem diferença estatística em relação a outros modelos multivariados. A classe SE é a que possui os melhores resultados de A.P. com limite inferior do IC acima de 0,75 para todos os modelos, indicando uma boa discriminabilidade desse alvo em relação às outras classes. Observam-se sobreposições na maioria dos modelos, o que revela similaridades estatísticas de desempenho nessa métrica. SE é a classe que apresenta menor diferença significativa entre os modelos utilizados, ou seja, é bem classificada em todos os modelos.

Figura 8 – Modas das 1000 classificações realizadas em um trecho da imagem PALSAR/ALOS, representando a classe mais frequente atribuída pelos modelos em cada *pixel*.



Fonte: Os autores (2025).

Figura 9 – Desempenho por modelo: médias e intervalo de credibilidade de 95% das métricas – Acurácia Global (preto), do Produtor (vermelho) e do Usuário (azul). Entre parênteses, “I” e “A” indicam dados em intensidade e amplitude, respectivamente. Gauss. = Gaussiano; biv. = bivariado.



Fonte: Os autores (2025).

No que diz respeito às classes florestais, F apresenta valores de A.P. abaixo de 0,5 para todos os modelos, o que expressa dificuldade em ser corretamente classificada. Para essa métrica em específico, observam-se largos ICs, principalmente para os modelos Gama HV, Gaussiano HH em intensidade e Gaussiano HH em amplitude. Nesses modelos, o limite inferior do IC é aproximadamente 0 (zero) para a A.P., enquanto os limites superiores estão próximos de 0,5. Esses resultados indicam a dificuldade dos modelos em predizerem corretamente a classe F e a forte dependência do conjunto amostral escolhido para gerar a classificação com valores melhores de acurácia. Por fim, para a classe VS não existe nenhum modelo onde o IC da A.P. não se sobrepõe ao de outro modelo. Assim como para F, alguns modelos apresentam grande variabilidade de valores, como, por exemplo, os modelos Gama HV, Gaussiano HH em intensidade e amplitude. A maior média é encontrada no modelo Gaussiano HV em amplitude (A.P. = 0,70). Para Acurácia do Usuário, observam-se ICs estreitos e médias abaixo de 0,5 para todos os modelos. Esses baixos valores indicam um maior erro de inclusão nessa classe, em que amostras de outras classes estão sendo atribuídas como VS em todos os modelos.

5 DISCUSSÕES

De modo geral, nenhum dos modelos avaliados apresentou desempenho superior em todas as métricas de avaliação – Acurácias Global, do Produtor e do Usuário. Embora diferenças de desempenho tenham sido observadas em nível de classe, nenhum modelo apresentou superioridade generalizada, tanto no cenário univariado quanto no bivariado.

Quanto ao pré-processamento das imagens SAR, as etapas de *multilooking*, filtragem *speckle*, *deskewing* e correção do terreno interferem nas propriedades estatísticas dos dados e, consequentemente, no número equivalente de *looks*. Essas operações tendem a aproximar a distribuição das amostras das classes da normalidade, o que pode justificar o desempenho similar dos modelos específicos para imagens SAR com os modelos Gaussianos. Cabe salientar, contudo, que tais operações de pré-processamento são executadas para tornar as imagens aptas a uso e, por essa razão, é comum que essas etapas sejam aplicadas de forma sistemática antes do processo de classificação. Nesse contexto, deve-se considerar o equilíbrio entre o grau de pré-processamento necessário para viabilizar uma aplicação e a aderência das amostras às distribuições estatísticas adotadas no classificador de Máxima Verossimilhança.

Por exemplo, Negri et al. (2016a) avaliaram o uso de imagens SAR banda-L polarimétricos para o mapeamento de classes de cobertura na Amazônia brasileira. Para tal, imagens PALSAR/ALOS foram classificadas com o classificador Maxver com a incorporação do algoritmo de regularização *Iterated Conditional Modes* (ICM) (Besag, 1986) sem a aplicação prévia de filtros de redução de *speckle* para preservar as propriedades estatísticas da imagem. Para dados univariados e dados totalmente polarizados, os autores realizaram testes de aderências às várias distribuições estatísticas próprias de imagens SAR, selecionando aquela em que as amostras se adequavam melhor.

Em contexto distinto, Reis et al. (2020a) avaliaram o impacto da utilização de dados SAR em estudos de detecção de mudanças na cobertura da terra em situações em que não há disponibilidade de imagens ópticas para uma ou mais datas, testando a aderência de amostras de imagens SAR, previamente submetidas a etapas típicas de pré-processamento (incluindo a filtragem), a diferentes distribuições estatísticas. Como resultado, verificaram que os dados apresentavam boa adequação à distribuição Gaussiana, a qual foi adotada por sua simplicidade de parametrização. Assim, considerando os resultados obtidos em nosso trabalho e os pré-processamentos aplicados, a adoção da distribuição Gaussiana na modelagem das classes mostra-se adequada, uma vez que não implica em perdas significativas nas métricas de acurácia – o que corrobora com os achados de Reis et al. (2020a).

Em relação às classes de cobertura, foram verificadas dificuldades de discriminação, principalmente no que diz respeito às classes florestais. Essa dificuldade também foi observada em outros trabalhos como Negri et al. (2016), Pereira et al. (2016) e Reis et al. (2020a) e era esperada nesse trabalho em função da análise exploratória das amostras e do dendrograma apresentado. Nos trabalhos citados, os autores encontraram melhores resultados com a utilização de legendas de cobertura menos detalhadas, agregando também as classes florestais. Uma estratégia que pode diminuir a confusão entre as classes e, consequentemente, melhorar as métricas de acurácia, consiste na utilização de séries temporais com inclusão de informações sobre

dinâmica das classes ao longo do tempo no processo de classificação, como realizado nos trabalhos de Dutra et al. (2023), Quevedo (2023) e Reis et al. (2020b).

Outro ponto importante é que, apesar de ser uma técnica tradicional de classificação, o Maxver apresenta vantagens práticas como simplicidade de parametrização, facilidade de implementação e baixo custo computacional de treinamento. Além disso, as verossimilhanças dos *pixels* em relação às classes de interesse podem ser utilizadas como estimativas de probabilidade, o que possibilita sua integração em outros modelos paramétricos de classificação como o *Compound Maximum a Posteriori* (CMAP) (Queiroz, 2025; Reis, 2022; Reis et al., 2020b). Essas vantagens explicam sua ampla difusão, estando disponível em *softwares* proprietários de processamento de imagens de Sensoriamento Remoto, como ENVI (Harris Geospatial), ERDAS IMAGINE (Hexagon Geospatial) e ArcGIS Pro (Esri), bem como em soluções livres, como o SNAP/STEP desenvolvido pela ESA, o Orfeo *Toolbox* (OTB) e o *Semi-Automatic Classification Plugin* (SCP) do QGIS. Nessas plataformas, o Maxver é implementado sob a hipótese de distribuição Gaussiana Multivariada. Assim, dados os resultados apresentados, observa-se que com um *workflow* que inclui processamentos como *multilooking* e filtragem de *speckle*, a modelagem Gaussiana pode ser empregada sem grandes perdas em relação à acurácia global das classificações, fazendo com que a implementação de distribuições mais complexas não fosse necessária, como a desenvolvida para um par de imagens em intensidade.

Por fim, como contribuição secundária, visando à reprodutibilidade e a utilização das distribuições estatísticas empregadas neste estudo, foi desenvolvido um conjunto de ferramentas em linguagem R, incluindo funções auxiliares relevantes para as análises realizadas. Os algoritmos estão disponibilizados em repositório público no GitHub (<https://github.com/vdlucas-queiroz/imageTools>).

6 CONCLUSÕES

Esta pesquisa avaliou a classificação da cobertura da terra utilizando dados SAR modelados por diferentes distribuições estatísticas. O uso das distribuições Gama e Par de imagens intensidade na classificação de imagens SAR foi avaliado em comparação com a distribuição Gaussiana. De modo geral, verificou-se que os modelos baseados em duas polarizações apresentaram valores de Acurácia Global (A.G.) superiores aos modelos que utilizaram apenas uma polarização. Além disso, o desempenho dos modelos Gaussianos bivariados tanto em intensidade como em amplitude são semelhantes aos do modelo Par de imagens intensidade (a 5% de significância). As sobreposições dos intervalos de credibilidade para A.G. indicam que modelos com mesma dimensionalidade e polarizações utilizadas possuem desempenhos globais de classificação similares. Assim, a depender do problema de classificação, onde sabidamente se aplicam as etapas de pré-processamento que envolvem médias e podem aproximar os dados de uma distribuição Gaussiana, é possível utilizá-la em substituição às distribuições específicas para dados SAR, preservando o desempenho global da classificação.

Os piores resultados de Acurácia Global foram observados para os modelos que utilizaram apenas a polarização HH, sendo o pior deles o modelo Gama HH. Além disso, o desempenho global de classificação utilizando apenas a polarização HV para os modelos Gama e Gaussiano em amplitude pode ser considerado semelhante ao do modelo Gaussiano bivariado em intensidade. No que diz respeito ao desempenho de classificação avaliado por classe, não foi verificado um modelo que se destacasse nas métricas de Acurácia do Produtor e do Usuário. Os resultados mostram que a classe Solo Exposto Agrícola é bem classificada em todos os modelos, contudo, a dissimilaridade entre as classes Floresta e Vegetação Secundária é baixa para imagens SAR, independentemente das distribuições estatísticas aplicadas. Sugere-se, então, a agregação das classes florestais na legenda para se obter menos confusões na classificação de cobertura.

Trabalhos futuros incluem a utilização de outras distribuições estatísticas adequadas para imagens SAR, a comparação de resultados com abordagens de classificação não-paramétricas como o *Random Forest*, a comparação de classificações com etapas de pré-processamento diferentes, observando seu impacto na aderência das amostras das classes e na classificação propriamente dita e o controle do número de *looks* em um range pré-estabelecido, verificando seu efeito nos resultados obtidos.

REFERÊNCIAS

- Almeida, C. A., Coutinho, A. C., Esquerdo, J. C. D. M., Adami, M., Venturieri, A., Diniz, C. G., Dessay, N., Durieux, L., & Gomes, A. R. (2016). *High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data*. <https://doi.org/10.1590/1809-4392201505504>
- Anfinson, S. N., Doulgeris, A. P., & Eltoft, T. (2009). Estimation of the equivalent number of looks in polarimetric Synthetic Aperture Radar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 47(11), 3795–3809. <https://doi.org/10.1109/TGRS.2009.2019269>
- Besag, J. (1986). On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48, 259–302.
- Brasil. (1974). Dispõe sobre a criação da Floresta Nacional do Tapajós e dá outras providências [Disponível em: <https://legislacao.presidencia.gov.br/atos/?tipo=DEC&numero=73684&ano=1974&ato=5c3oXQE1EMnRVT95b>. Acesso em: 25 out. 2025.].
- Correia, A. H. (1999). *Projeto, desenvolvimento e avaliação de classificadores estatísticos pontuais e contextuais para imagens SAR polarimétricas* [Mestrado em Sensoriamento Remoto]. Instituto Nacional de Pesquisas Espaciais (INPE). <http://urlib.net/sid.inpe.br/deise/1999/09.14.14.45>
- Dutra, L. V., Rennó, C. D., Reis, M. S., & Gamba, P. (2023). A Generative Method for Simultaneous Classification of Remote Sensing Time Series Data Using and Ensemble of Decision Tree Classifiers. *Anais...*, 179–182.
- European Space Agency. (2007). *Information on ALOS PALSAR products for ADEN users*.
- Frery, A. C., Correa, A., Renno, C. D., Freitas, C. C., Jacobo-Berlles, J., Vasconcellos, K. L. P., Mejail, M., & Sant'anna, S. J. S. (1999). Models for Synthetic Aperture Radar image analysis. *Resenhas do Instituto de Matematica e Estatística da Universidade de São Paulo*, 4(1), 45–77.
- Frery, A. C., Correia, A. H., & Freitas, C. C. (2007). Classifying multifrequency fully polarimetric imagery with multiple sources of statistical evidence and contextual information. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10), 3098–3109. <https://doi.org/10.1109/TGRS.2007.903828>
- Frery, A. C., Freitas, C., Sant'anna, S. J. S., & Rennó, C. D. (1999). Statistical properties of SAR data and their consequences. *United Nations Programme on Space Applications*, 10, 53.
- Frery, A. C., Vieira, P. R., Sant'anna, S. J. S., Rennó, C. D., et al. (1997). A user-friendly system for Synthetic Aperture Radar image classification based on grayscale distributional properties and context. *Proceedings X Brazilian Symposium on Computer Graphics and Image Processing*, 211–218.
- Gao, G. (2010). Statistical modeling of SAR images: A survey. *Sensors*, 10(1), 775–795.
- Garg, A., & Singh, D. (2018). Development of an efficient contextual algorithm for discrimination of tall vegetation and urban for PALSAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 56(6), 3413–3420.
- Jensen, J. R. (2014). *Remote Sensing of the environment: an earth resource perspective* (Vol. 2). Pearson Education Limited.
- Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*, 4, 83–91.
- Kwak, S., & Kim, J. H. (2017). Central limit theorem: the cornerstone of modern statistics. *Korean Journal of Anesthesiology*, 70, 144–156. <https://doi.org/10.4097/kjae.2017.70.2.144>

- Lambin, E. F., Geist, H., & Rindfuss, R. R. (2006). Introduction: local processes with global impacts. *Land-use and land-cover change: Local processes and global impacts*, 1–8.
- Lee, J. S., Hoppel, K. W., Mango, S. A., & Miller, A. R. (1994). Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 32(5), 1017–1028. <https://doi.org/10.1109/36.312890>
- Lee, J., & Pottier, E. (2017). *Polarimetric radar imaging: from basics to applications*. CRC press.
- Mahapatra, D. K., Frery, A. C., Pradhan, B. B., & Roy, L. P. (2024). A Review on parametric and semiparametric distributions in characterizing Synthetic Aperture Radar clutter data. *IEEE Access*, 12, 83340–83362.
- Martins, V. S., Novo, E. M., Lyapustin, A., Aragão, L. E., Freitas, S. R., & Barbosa, C. C. (2018). Seasonal and interannual assessment of cloud cover and atmospheric constituents across the Amazon (2000–2015): insights for remote sensing and climate analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 309–327.
- Meyer, F. (2019). Spaceborn Synthetic Aperture Radar: principles data access and basic process techniques [Astrophysics and Space Science Proceedings]. Em A. I. Flores-Anderson, K. E. Herndon, R. B. Thapa & E. Cherrington (Ed.), *The SAR handbook: comprehensive methodologies for forest monitoring and biomass estimation*. (Pp. 21–64). National Space Science; Technology Center.
- Negri, R. G., Dutra, L. V., Freitas, C. C., & Lu, D. (2016a). Exploring the Capability of ALOS PALSAR L-Band Fully Polarimetric Data for Land Cover Classification in Tropical Environments. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(12), 5369–5384.
- Negri, R. G., Dutra, L. V., Sant’Anna, S. J. S., & Lu, D. (2016). Examining region-based methods for land cover classification using stochastic distances. *International Journal of Remote Sensing*, 37(8), 1902–1921.
- Oliver, C., & Quegan, S. (2004). *Understanding Synthetic Aperture Radar images*. SciTech Publishing.
- Pereira, L. O., Freitas, C. C., Sant’Anna, S. J. S., & Reis, M. S. (2018). Evaluation of optical and radar images integration methods for LULC classification in Amazon region. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(9), 3062–3074.
- Pereira, L. O., Freitas, C. C., Sant’Anna, S. J. S., & Reis, M. S. (2016). ALOS/PALSAR data evaluation for land use and land cover mapping in the Amazon region. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(12), 5413–5423. <https://doi.org/10.1109/JSTARS.2016.2622481>
- Queiroz, V. D. B. (2025). *Classificação de trajetórias de cobertura da terra utilizando uma abordagem de máximo a posteriori composto em imagens de multissensores* [diss. de mestr., Instituto Nacional de Pesquisas Espaciais (INPE)]. <http://urlib.net/ibi/8JMKD3MGP3W34T/4BRJ5G5>
- Quevedo, R. P. (2023). *Do land use and land cover and spatial heterogeneity influence on landslide occurrence? An analysis of susceptibility* [Doutorado em Sensoriamento Remoto]. Instituto Nacional de Pesquisas Espaciais (INPE). <http://urlib.net/ibi/8JMKD3MGP3W34T/49AT9R8>
- Reis, M. S. (2022). *Detection and analysis of forest regeneration trajectories in the lower Tapajós region* [Tese (Doutorado em Ciência do Sistema Terrestre)]. Instituto Nacional de Pesquisas Espaciais (INPE). <http://urlib.net/ibi/8JMKD3MGP3W34T/47E2TRB>
- Reis, M. S., Barros, L. S., Rodrigues Neto, M. R., Moraes, D. R. V., Moreira, N. A. P., Alves, G. M. R., Adorno, B. V., Messias, C. G., Dutra, L. V., Rennó, C. D., Sant’Anna, S. J. S., & Escada, M. I. S. (2024). Assessing interpreter’s disagreements in land cover reference data collection from

- historical Landsat time series in Amazon. *International Journal of Remote Sensing*, 45(15), 5192–5223. <https://doi.org/10.1080/01431161.2024.2373340>
- Reis, M. S., Dutra, L. V., & Escada, M. I. S. (2017). Simultaneous multi-source and multi-temporal land cover classification using a compound maximum likelihood classifier. *GEOINFO*, 18, 74–85.
- Reis, M. S., Dutra, L. V., Escada, M. I. S., & Sant'anna, S. J. S. (2020b). Avoiding invalid transitions in land cover trajectory classification with a Compound Maximum a Posteriori approach. *IEEE Access*, 8, 98787–98799. <https://doi.org/10.1109/ACCESS.2020.2997019>
- Reis, M. S., Dutra, L. V. D., Sant'Anna, S. J. S., & Escada, M. I. S. (2020a). Multi-source change detection with PALSAR data in the Southern of Pará state in the Brazilian Amazon. *International Journal of Applied Earth Observation and Geoinformation*, 84, 101945. <https://doi.org/https://doi.org/10.1016/j.jag.2019.101945>
- Reis, M. S., Escada, M. I. S., Dutra, L. V., Sant'Anna, S. J. S., & Vogt, N. D. (2018). Towards a reproducible LULC hierarchical class legend for use in the southwest of Pará state, Brazil: A comparison with Remote Sensing data-driven hierarchies. *Land*, 7(2). <https://doi.org/10.3390/land7020065>
- Richards, J. A. (2009). *Remote Sensing with imaging Radar* (Vol. 1). Springer. <https://doi.org/10.1007/978-3-642-02020-9>
- Richards, J. A. (2022). *Remote sensing digital image analysis: an introduction* (Vol. 6). Springer.
- Sagrillo, M., Guerra, R. R., Bayer, F. M., & Machado, R. (2022). A new probability distribution for SAR image modeling. *Remote Sensing*, 14(12).
- Sant'Anna, S. J. S., Braga, B. C., Oliveira, J. M., Oliveira, M. A. F., Reis, M. S., Moreira, N. A. P., & Albuquerque, P. C. G. (2016). Field data from the Tapajós region - August-September of 2016.
- Schowengerdt, A. R. (2007). *Remote sensing: models and methods for image processing* (3^a ed.). Academic Press.
- Servello, E. L. (2010). Dados polarimétricos do RADARSAT-2 (banda C) na discriminação de uso e cobertura da terra na região da Floresta Nacional do Tapajós.
- Shimada, M., Isoguchi, O., Tadono, T., & Isono, K. (2009). PALSAR radiometric and geometric calibration. *IEEE Transactions on Geoscience and Remote Sensing*, 47(12), 3915–3932. <https://doi.org/10.1109/TGRS.2009.2023909>
- Smirnov, N. (1939). On the estimation of discrepancy between empirical curves of distribution for two independent samples. *Bulletin Mathématique de L'Université de Moscow*, 2(2), 3–11.
- Theodoridis, S., & Koutroumbas, K. (2009). *Pattern recognition* (4^a ed.). Academic Press.
- Torres, L. J. T. M. (2012). *Um novo algoritmo para filtragem de speckle em imagens SAR de intensidade baseado em distâncias estocásticas* [Mestrado em Modelagem Computacional de Conhecimento]. Universidade Federal de Alagoas (UFAL). <http://www.repositorio.ufal.br/jspui/handle/riufal/1829>
- Turner, B., & Meyer, W. (1994). Global land-use and land-cover change: an overview. Em W. B. Meyer & B. L. Turner (Ed.), *Changes in land use and land cover: a global perspective*. Cambridge University Press.
- Veci, L. (2015). *ALOS PALSAR Orthorectification Tutorial*. Array Systems Computing Inc. / ESA STEP. <https://step.esa.int/docs/tutorials/ALOS%20PALSAR%20Orthorectification%20Tutorial.pdf>
- Veloso, H. P., RANGEL FILHO, A. L. R., & Lima, J. C. A. (1991). *Classificação da vegetação brasileira, adaptada a um sistema universal*. IBGE.
- Ward Jr., J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236–244. <https://doi.org/10.1080/01621459.1963.10500845>

- Woodhouse, I. H. (2006). *Introduction to microwave Remote Sensing*. CRC press Taylor & Francis.
- Wright, P. A., Meadows, P. J., Mack, G., Miranda, N., & Lavalley, M. (2009). ADEN ALOS PALSAR product verification. *Proceedings...* <https://earth.esa.int/eogateway/documents/20142/37627/ALOS-PI-Symposium-PALSAR-Verification.pdf>
- Wulder, M. A., Hilker, T., White, J. C., Coops, N. C., Masek, J. G., Pflugmacher, D., & Crevier, Y. (2015). Virtual constellations for global terrestrial monitoring. *Remote Sensing of Environment*, 170, 62–76.
- Yanasse, C. C. F. (1991). *Statistical analysis of Synthetic Aperture Radar images and its applications to system analysis and change detection*. [tese de dout., University of Sheffield].
- Yanasse, C. S. F., Sant’Anna, S. J. S., Frery, A. C., Rennó, C. D., Soares, J. V., & Luckman, A. J. (1997). Exploratory study of the relationship between tropical forest regeneration stages and SIR-C L and C data. *Remote Sensing of Environment*, 59(2), 180–190. [https://doi.org/10.1016/s0034-4257\(96\)00149-6](https://doi.org/10.1016/s0034-4257(96)00149-6)
- Yue, D. X., Xu, F., Frery, A. C., & Jin, Y. Q. (2021). SAR image statistical modeling part I: single-pixel statistical models. *IEEE Geoscience and Remote Sensing Magazine*, 9(1), 82–114.

Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Código de Financiamento 001) pela bolsa de estudos (processo nº 88887.668994/2022-00), bem como ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), projeto nº 422354/2023-6, “Monitoramento e Avisos de Mudanças de Cobertura da Terra nos Biomas Brasileiros – Treinamento e Semi-Automação do Programa BiomasBR”, bolsa nº 313799/2025-3 pelo apoio. Agradecem, ainda, os valiosos comentários recebidos durante o processo de revisão deste manuscrito.

Contribuição dos Autores

Conceptualização: V.D.L.B.Q., M.S.R e S.J.S.S. Curadoria dos dados: V.D.L.B.Q. e M.S.R. Análise formal: V.D.L.B.Q., M.S.R e S.J.S.S. Investigação: V.D.L.B.Q. Metodologia: V.D.L.B.Q., M.S.R, e S.J.S.S. Administração do projeto: V.D.L.B.Q. e M.S.R. Geração de resultados: V.D.L.B.Q. Supervisão: M.S.R e S.J.S.S. Validação: V.D.L.B.Q., M.S.R, e S.J.S.S. Visualização: V.D.L.B.Q. Redação–minuta inicial: V.D.L.B.Q. Redação–revisão e edição: V.D.L.B.Q, M.S.R e S.J.S.S.

Conflitos de Interesse

Não há conflitos de interesse.

Biografia do autor principal



Vinícius D’Lucas Bezerra e Queiroz se formou em Engenharia Cartográfica e de Agrimensura pela Universidade Federal de Pernambuco, *campus* Recife, em 2021. Concluiu em 2024 o mestrado em Sensoriamento Remoto pelo Instituto Nacional de Pesquisas Espaciais (INPE), em São José dos Campos - SP. Atualmente é bolsista de treinamento técnico na Universidade Federal de São Paulo (UNIFESP), atuando nas áreas de aprendizado de máquina e ciência cidadã. Seus interesses principais incluem técnicas de processamento de imagens ópticas e de radares de abertura sintética e o desenvolvimento de algoritmos de classificação.



Esta obra está licenciada com uma Licença [Creative Commons Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) – CC BY. Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original.