



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA (CCEN)  
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA (PPGE)

Joselito Elias de Araújo

**Modelagem Estatística e Redução de Ruído em Dados de Interferometria SAR**

Recife

2025

Joselito Elias de Araújo

## **Modelagem Estatística e Redução de Ruído em Dados de Interferometria SAR**

Tese apresentada ao Programa de Pós-graduação em Estatística do Centro de Ciências Exatas e da Natureza (CCEN) da Universidade Federal de Pernambuco (UFPE) como requisito para obtenção do título de Doutor em Estatística.

**Área de Concentração:** Probabilidade e Estatística

**Orientador:** Prof. Dr. Alejandro César Frery Orgambide

**Coorientador:** Prof. Dr. Anderson Adaime de Borba

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Araújo, Joselito Elias de.

Modelagem estatística e redução de ruído em dados de interferometria SAR / Joselito Elias de Araújo. - Recife, 2025. 105f.: il.

Tese (Doutorado) - Universidade Federal de Pernambuco, Centro de Ciências Exatas e da Natureza, Programa de Pós-Graduação em Estatística, 2025.

Orientação: Alejandro César Frery Orgambide.

Coorientação: Anderson Adaime de Borba.

Inclui referências e apêndices.

1. Imagens SAR; 2. Interferometria SAR; 3. Polarimetria SAR; 4. Modelos Estatísticos; 5. Diferença de Fase; 6. Filtragem. I. Orgambide, Alejandro César Frery. II. Borba, Anderson Adaime de. III. Título.

UFPE-Biblioteca Central

**JOSELITO ELIAS DE ARAÚJO**  
**MODELAGEM ESTATÍSTICA E REDUÇÃO DE RUÍDO EM DADOS DE INTERFEROMETRIA SAR**

Tese apresentada ao Programa de Pós Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Estatística.

**Aprovado em:** 25 de fevereiro de 2025.

**BANCA EXAMINADORA**

Prof. Dr. Alejandro César Frery Orgambide  
Presidente (Orientador), Victoria University  
of Wellington

Prof. Dr. Anderson Adaime de Borba  
Coorientador, Mackenzie

Prof. Dr. Renato Machado  
Examinador Externo à Instituição, ITA

Profa. Dra. Fátima N. Sombra de Medeiros  
Externa ao Programa, UFPE

Prof. Dr. Abraão D. C. Nascimento  
Examinador Interno, UFPE

Prof. Dr. Getúlio José do Amaral  
Examinador Interno, UFPE

Dedico este trabalho a minha família.

## AGRADECIMENTOS

Agradeço, primeiramente a Deus, pela vida e pela força que tem me dado todos os dias, assim como às pessoas do meu convívio que acreditaram e contribuíram, mesmo que indiretamente, para a conclusão deste curso.

Aos meus pais, José Araújo Filho e Maria Elias de Araújo, pelo amor e pela paciência que têm me dedicado ao longo desse tempo. Agradeço por terem feito o possível e o impossível para me oferecer a oportunidade de estudar, acreditando e respeitando minhas decisões, nunca permitindo que as dificuldades acabassem com os meus sonhos.

À minha esposa, Luzineide do Nascimento Silva, e aos meus filhos, Arthur Nascimento Araújo e Lorena Nascimento Araújo, por terem compartilhado comigo todas as angústias e felicidades, acompanhando de perto cada passo. Agradeço pelo amor, carinho e apoio depositados, além da companhia ao longo de todos esses anos.

Aos meus orientadores, professores Alejandro César Frery Orgambide e Anderson Adaime de Borba, pela paciência, dedicação, compromisso e ensinamentos que possibilitaram a realização deste trabalho.

Agradeço aos meus amigos e companheiros de curso: Edison Fernando, José Iraponil, Jerfson Bruno, José Jairo, Tatiane, Ivangllys, Ernando Campos, Maria Júlia, Alisson de Lima, Vinícius, Eduardo e Rosa Janeth, por contribuírem com a minha aprendizagem e por palavras de incentivo.

À Universidade Federal de Pernambuco (UFPE), em especial a todos os professores do Departamento de Estatística desta instituição, que desempenham seu trabalho com compromisso e dedicação.

Por fim, agradeço à Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) pelo apoio financeiro.

## RESUMO

Sistemas de sensoriamento remoto requerem imagens de alta resolução para capturar e fornecer informações precisas das áreas monitoradas. Sob essa perspectiva, os sistemas SAR (Synthetic Aperture Radar) oferecem recursos para obter essas imagens. No entanto, a qualidade das imagens é frequentemente comprometida pelo fenômeno conhecido como speckle, uma interferência que gera uma textura granulada, dificultando a interpretação dos dados. Este problema é particularmente relevante em sistemas SAR, que utilizam a coerência de sinais para obter imagens detalhadas e confiáveis. Além disso, o PolSAR (Polarimetric SAR) e o InSAR (Interferometric SAR) ampliam as possibilidades de análise, fornecendo informações polarimétricas e interferométricas, e diversos métodos, incluindo a aplicação de modelos estatísticos, são utilizados para extrair informações valiosas dessas observações. Todavia, esses métodos também enfrentam desafios relacionados ao ruído e à precisão na modelagem dos dados. Diante desse contexto, essa pesquisa busca responder à seguinte questão: como desenvolver algoritmos de filtragem eficientes para reduzir o ruído de fase, preservando informações e melhorando a qualidade das imagens SAR/InSAR? O objetivo principal é aprofundar o conhecimento em análise estatística de dados circulares, com destaque na avaliação da qualidade de imagens SAR/InSAR. Para alcançar esse objetivo, revisamos a literatura sobre modelos estatísticos para dados circulares, enfatizando modelos físicos e empíricos. Disponibilizando essas soluções por meio dos pacotes `FDDPhase` e `Filters` na plataforma R. Além disso, utilizamos a técnica de MLE (Maximum Likelihood Estimation) para estimar os parâmetros de modelos, validando os resultados tanto com dados sintéticos quanto com dados SAR. A eficácia dos métodos propostos foi avaliada por meio de métricas como: SSIM (Structural Similarity Index), RMSE (Root Mean Square Error), média da diferença, variância da diferença, tempo de execução e número de resíduos. Os resultados obtidos nesta tese mostram que os métodos propostos reduzem significativamente o speckle e superam o desempenho computacional do Lee Refinado.

**Palavras-chaves:** Imagens SAR. Interferometria SAR. Polarimetria SAR. Modelos Estatísticos. Diferença de Fase. Filtragem.

## ABSTRACT

Remote sensing systems require high-resolution images to capture and provide precise information about the monitored areas. In this regard, Synthetic Aperture Radar (SAR) systems offer the necessary tools to acquire such images. However, the quality of these images is often compromised by a phenomenon known as speckle, an interference that generates a grainy texture, making data interpretation more difficult. This issue is particularly relevant in SAR systems, which rely on signal coherence to obtain detailed and reliable images. Furthermore, Polarimetric SAR (PolSAR) and Interferometric SAR (InSAR) expand the possibilities for analysis by providing polarimetric and interferometric information, with various methods, including applying statistical models, used to extract valuable insights from these observations. However, these methods also face challenges related to noise and data modeling accuracy. In this context, this research aims to answer the following question: how can efficient filtering algorithms be developed to reduce phase noise while preserving information and improving the quality of SAR/InSAR images? The primary objective is to deepen the understanding of statistical analysis of circular data, with a focus on evaluating the quality of SAR/InSAR images. To achieve this goal, we reviewed the literature on statistical models for circular data, highlighting both physical and empirical models. We made these solutions available through the `FDDPhase` and `Filters` packages on the R platform. We used the Maximum Likelihood Estimation (MLE) technique to estimate the model parameters, validating the results with synthetic and SAR data. The effectiveness of the proposed methods was evaluated using metrics such as SSIM (Structural Similarity Index), RMSE (Root Mean Square Error), mean difference, variance of the difference, execution time, and the number of residues. The results obtained in this thesis show that the proposed methods significantly reduce speckle and outperform the computational performance of the Refined Lee filter.

**Keywords:** SAR images. SAR interferometry. SAR polarimetry. Statistical models. Phase difference. Filtering.

## LISTA DE FIGURAS

Figura 1 – Geometria básica de um sistema de imageamento por Radar de Abertura Sintética . . . . .	23
Figura 2 – Direções de propagação do campo elétrico: (a) polarização vertical (V) e (b) polarização horizontal (H) . . . . .	24
Figura 3 – Divisão da abertura sintética em segmentos ou visadas . . . . .	26
Figura 4 – Interferometria de passagem única e dupla . . . . .	28
Figura 5 – Geometria básica de um sistema InSAR . . . . .	29
Figura 6 – Diferença de fase ou fase interferométrica . . . . .	30
Figura 7 – Representação de duas imagens complexas co-registradas . . . . .	31
Figura 8 – Múltiplas visadas para distribuição de diferença de fase: Jong-Sen Lee . . . .	46
Figura 9 – Múltiplas visadas para distribuição de diferença de fase: Christoph H. Gierull	48
Figura 10 – Modelos estatísticos . . . . .	53
Figura 11 – Relações entre modelos estatísticos . . . . .	54
Figura 12 – Operação de convolução . . . . .	55
Figura 13 – Interferograma exibindo ciclos de franjas coloridas . . . . .	56
Figura 14 – Janelas direcionais para o filtro Lee InSAR Refinado . . . . .	58
Figura 15 – Janelas direcionais para os filtros propostos . . . . .	62
Figura 16 – Ajustes entre os modelos físicos . . . . .	71
Figura 17 – Erro Absoluto entre os modelos em (3.20) e (3.21): $L = 3/L = 3,49$ (vermelho), $L = 5/L = 5,49$ (azul), $L = 7/L = 7,49$ (verde) e $L = 9/L = 9,49$ (roxo). . . . .	72
Figura 18 – (a) Fase original(livre de ruído) (b) Fase ruidosa . . . . .	73
Figura 19 – Antes e depois da filtragem: (a) Lee Refinado, (b) LInSARRFE, (c) TcN-filter, (d) TcCfilter. Fase desembrulhada após a aplicação do filtro: (e) Lee Refinado, (f) LInSARRFE, (g) TcNfilter, (h) TcCfilter . . . . .	73
Figura 20 – Imagem de diferença de fase ou interferograma . . . . .	76
Figura 21 – Antes e depois da filtragem: (a) Lee InSAR Refinado, (b) LInSARRFE, (c) TcNfilter, (d) TcCfilter. Fase desembrulhada após aplicação do filtro: (e) Lee InSAR Refinado, (f) LInSARRFE, (g) TcNfilter, (h) TcCfilter . . . .	76
Figura 22 – Imagem de diferença de fase ou interferograma . . . . .	77

Figura 23 – Comparação de resultados filtrados: (a) Lee InSAR Refinado, (b) LInSAR-RFE, (c) TcNfilter, (d) TcCfilter. Fase desembrulhada após aplicação do filtro: (e) Lee InSAR Refinado, (f) LInSARRFE, (g) TcNfilter, (h) TcCfilter	78
Figura 24 – Tempo Otimizado . . . . .	79
Figura 25 – Função de Bessel de primeira espécie: $J_\nu(x)$ . . . . .	92
Figura 26 – Função de Bessel modificada de primeira espécie: $I_\nu(x)$ . . . . .	94
Figura 27 – Relação entre coordenadas cartesianas e polares . . . . .	95
Figura 28 – Representação de uma distribuição circular contínua . . . . .	98

## LISTA DE TABELAS

Tabela 1 – Características e relações entre modelos estatísticos . . . . .	54
Tabela 2 – Distância entre modelos físicos truncando o número de looks . . . . .	70
Tabela 3 – Eficiência dos filtros para dados simulados . . . . .	74
Tabela 4 – Eficiência dos filtros para dados SAR – La Cumbre . . . . .	77
Tabela 5 – Eficiência dos filtros para dados SAR – Irã-Iraque . . . . .	79

## LISTA DE ABREVIATURAS E SIGLAS

<i>CT</i>	Cauchy Truncada
<i>DC</i>	Distribuição Cardioide
<i>DUC</i>	Distribuição Uniforme Circular
<i>NT</i>	Normal Truncada
<i>VAC</i>	Variável Aleatória Circular
<i>VM</i>	von Mises
<i>WN</i>	Cauchy Wrapped
<i>WN</i>	Normal Wrapped
<b>CDF</b>	Função de Distribuição Acumulada
<b>DWCE</b>	Distribuição Wishart Complexa Escalonada
<b>EDO</b>	Equação Diferencial Ordinária
<b>GHF</b>	Função Hipergeométrica de Gauss
<b>LInSARRFE</b>	Extensão do Filtro Lee InSAR Refinado
<b>PDF</b>	Função Densidade de Probabilidade
<b>ROI</b>	Região de Interesse
<b>TcCfilter</b>	Filtro Cauchy Truncada
<b>TcNfilter</b>	Filtro Normal Truncada

## LISTA DE SÍMBOLOS

$\psi$	Diferença de fase
$\rho_c$	Coeficiente de correlação complexo
$\psi_\xi$	Limite de fase
$E_V$	Campo elétrico vertical
$E_H$	Campo elétrico horizontal
$\underline{E}^i$	Onda incidente polarizada
$\underline{E}^s$	Onda dispersada
H	Polarização horizontal
V	Polarização vertical
X	Variável aleatória escalar
L	Número de looks
$\Theta$	Variável aleatória circular
$\mathbf{X}$	Matriz aleatória
$\mathbf{S}$	Matriz de espalhamento
$\Sigma$	Matriz de covariância
$\mathbf{Z}$	Matriz de covariância amostral

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
1.1	OBJETIVOS	18
1.2	PLATAFORMAS COMPUTACIONAIS	19
1.3	TRABALHOS PUBLICADOS	19
1.4	ORGANIZAÇÃO DO TRABALHO	20
<b>2</b>	<b>BASE TEÓRICA E OPERACIONAL DE SISTEMAS SAR</b>	<b>21</b>
2.1	SISTEMAS SAR/POLSAR E SUAS CARACTERÍSTICAS	21
2.2	RADARES SAR POLARIMÉTRICOS E POLARIZAÇÃO	22
2.3	RUÍDO SPECKLE	25
2.4	INTERFEROMETRIA SAR	27
<b>2.4.1</b>	<b>Processo de aquisição de imagens SAR</b>	<b>28</b>
<b>2.4.2</b>	<b>Diferença de fase ou fase interferométrica</b>	<b>29</b>
2.5	MODELAGEM PARA DADOS POLSAR MULTILOOK	31
<b>2.5.1</b>	<b>Estatísticas Gaussianas</b>	<b>36</b>
<b>2.5.2</b>	<b>Coefficiente de correlação complexo</b>	<b>37</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>39</b>
3.1	MODELOS PROBABILÍSTICOS PARA DADOS CIRCULARES	39
<b>3.1.1</b>	<b>Modelos circulares empíricos</b>	<b>39</b>
3.1.1.1	<i>Distribuição Uniforme Circular</i>	40
3.1.1.2	<i>Distribuição Cardioide</i>	41
3.1.1.3	<i>Distribuição normal wrapped</i>	41
3.1.1.4	<i>Distribuição de Cauchy wrapped</i>	42
3.1.1.5	<i>Distribuição de von Mises</i>	43
<b>3.1.2</b>	<b>Modelos físicos para diferença de fases</b>	<b>44</b>
3.2	MODELOS EMPÍRICOS TRUNCADOS	48
3.3	FUNÇÃO DE LOG-VEROSSIMILHANÇA	50
<b>3.3.1</b>	<b>Função de log-verossimilhança do modelo definido em (3.20)</b>	<b>50</b>
<b>3.3.2</b>	<b>Função de log-verossimilhança do modelo definido em (3.21)</b>	<b>51</b>
3.4	RELAÇÕES ENTRE MODELOS ESTATÍSTICOS	52
3.5	PROCESSO DE FILTRAGEM	54

3.6	FILTROS DE FASE PARA DADOS INSAR . . . . .	56
<b>3.6.1</b>	<b>Filtro adaptativo Lee InSAR . . . . .</b>	<b>58</b>
<b>3.6.2</b>	<b>Filtro Lee InSAR Refinado . . . . .</b>	<b>59</b>
<b>3.6.3</b>	<b>Algoritmos de Filtragem Propostos . . . . .</b>	<b>61</b>
3.7	DIVERGÊNCIA ESTATÍSTICA . . . . .	64
<b>3.7.1</b>	<b>Distâncias Estocásticas . . . . .</b>	<b>64</b>
3.8	PROCEDIMENTOS E MÉTODOS . . . . .	65
<b>3.8.1</b>	<b>Levantamento e Modelagem Estatística . . . . .</b>	<b>65</b>
<b>3.8.2</b>	<b>Aplicação dos Modelos e Filtros . . . . .</b>	<b>65</b>
<b>3.8.3</b>	<b>Avaliação do Desempenho . . . . .</b>	<b>66</b>
<b>4</b>	<b>RESULTADOS . . . . .</b>	<b>69</b>
4.1	DISTÂNCIA ENTRE MODELOS . . . . .	69
<b>4.1.1</b>	<b>Distância entre os modelos (3.20) e (3.21) . . . . .</b>	<b>69</b>
4.2	FILTRAGEM DE DADOS SIMULADOS . . . . .	72
4.3	FILTRAGEM DE DADOS INSAR . . . . .	75
<b>4.3.1</b>	<b>La Cumbre . . . . .</b>	<b>75</b>
<b>4.3.2</b>	<b>Fronteira Irã-Iraque . . . . .</b>	<b>77</b>
4.4	DESEMPENHO TEMPORAL DOS FILTROS PROPOSTOS EM COMPA- RAÇÃO AO LEE REFINADO . . . . .	79
<b>5</b>	<b>CONSIDERAÇÕES FINAIS . . . . .</b>	<b>81</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>83</b>
	<b>APÊNDICE A – FUNÇÕES DE BESSEL . . . . .</b>	<b>88</b>
	<b>APÊNDICE B – ESTATÍSTICA CIRCULAR . . . . .</b>	<b>95</b>

## 1 INTRODUÇÃO

Os radares foram desenvolvidos inicialmente para fins militares, durante a Segunda Guerra Mundial e posteriormente, a partir da década de 70, passaram a ser largamente utilizados também, para fins civis. O primeiro Radar de Abertura Sintética (SAR – Synthetic Aperture Radar) foi lançado pela Administração Nacional de Aeronáutica e Espaço (NASA – National Aeronautics and Space Administration) em 1978. O SAR surgiu para resolver um problema encontrado nos Radares de Abertura Real (RAR – Real Aperture Radar, SLAR – Side Looking Airbone Radar), pois a resolução em azimute das imagens era inversamente proporcional ao comprimento da antena, ou seja, para se obter uma imagem de alta qualidade era necessário uma antena com dezenas ou até centenas de metros de comprimento, o que era impossível de se obter em plataformas orbitais.

O SAR é um sistema coerente, ou seja, preserva a fase do sinal durante todo o processo de aquisição e processamento dos dados. O SAR representa um modo engenhoso de se superar o problema do tamanho físico da antena, sintetizando uma antena virtual por meio do registro acumulado dos sinais de retorno de cada ponto da antena, durante o período em que ela se desloca sobre uma dada região do terreno (NOVO, 2010).

Uma vez que o pulso tenha passado sobre um ponto do terreno, toda a informação de fase sobre aquele ponto é armazenada em uma matriz bidimensional (range e azimute). Toda a "história" de fase de todos os pontos da imagem é combinada numa série temporal que forma o dado sintético. Através de um processamento complexo, esta "assinatura de fase" de cada ponto é transformada em informação de azimute e range (NOVO, 2010).

Assim como o SAR, o Radar de Abertura Sintética Polarimétrica (PolSAR - Polarimetric Synthetic Aperture Radar) é um sistema de radar coerente que utiliza múltiplos canais de polarização, como vertical (V) e horizontal (H), para obter informações detalhadas sobre a área monitorada. Ao simular eletronicamente uma antena ou abertura extremamente grande, utilizando a trajetória de voo da plataforma, o PolSAR gera imagens com alta resolução. Sua principal característica é a capacidade de medir as propriedades polarimétricas dos retornos retroespalhados das ondas de radar, proporcionando valiosas informações em sensoriamento remoto.

O sensoriamento remoto, por sua vez, abrange uma ampla variedade de aplicações, como monitoramento ambiental, agricultura de precisão, análise de desastres naturais e mapeamento de recursos terrestres. Sistemas SAR são essenciais para gerar dados detalhados e atualizados

sobre a superfície terrestre, desempenhando um papel fundamental na coleta de informações para diversas áreas do conhecimento.

Desde a aquisição da primeira imagem SAR nos Estados Unidos, a análise de dados SAR promoveu diretamente o desenvolvimento de técnicas de modelagem estatística. No entanto, um problema inerente ao imageamento com iluminação coerente, como é o caso do SAR/PolSAR, é a presença de um padrão de interferência, denominado speckle. Como o speckle é um ruído multiplicativo, não é adequado utilizar a distribuição gaussiana na descrição dos dados de imagens PolSAR multilook (FRERY et al., 1997). Assim, a inclusão de recursos polarimétricos nas imagens possibilita a obtenção de informações mais precisas sobre a área observada, o que melhora a qualidade dos dados. No entanto, essa ampliação também traz desafios adicionais no processamento das imagens, sendo um deles a mitigação do ruído.

O speckle é um dos principais fatores que degradam a qualidade de imagens SAR. Existem dois métodos para diminuir o ruído speckle nas imagens SAR: a filtragem e o processamento multilook. Os filtros devem preservar o valor médio do retorno do radar (backscatter) e garantir que características importantes da imagem, como bordas e textura, não sejam comprometidas durante o processo de filtragem. Por esse motivo, as propostas de técnicas específicas para reduzir o speckle tiveram origem há mais de trinta anos (LEE, 1980) e (KUAN et al., 1987) e ainda permanecem efetivas (ARGENTI et al., 2013), (CHAO; CHEN; LEE, 2013), (MA et al., 2018) e (LI; CHEN; LEE, 2020).

A modelagem estatística é uma das abordagens mais frutíferas para a interpretação de imagens SAR (GAO, 2010). Ela abrange diversas áreas, como reconhecimento de padrões, processamento de imagens, análise de sinais, teoria da probabilidade e estudo das características de espalhamento eletromagnético dos alvos, entre outras. Nesse contexto, a modelagem estatística desempenha um papel fundamental na análise de imagens interferométricas, uma técnica que começa com a aquisição de duas imagens SAR complexas da mesma área de interesse, obtidas em momentos ou ângulos ligeiramente diferentes.

Imagens SAR/InSAR, que registram informações de amplitude e fase, devem ser co-registradas para garantir que os pixels correspondentes representem o mesmo alvo no terreno. Após o co-registro, a diferença de fase é obtida ao calcular o produto entre os valores complexos dos pixels de uma imagem e o conjugado transposto dos pixels da outra. O resultado desse cálculo é uma nova imagem complexa, chamada interferograma, na qual cada pixel armazena a fase interferométrica, que está diretamente relacionada à elevação do terreno ou às mudanças em sua superfície, dependendo da aplicação (topografia ou deformação).

Segundo Wang et al. (2024), com os rápidos avanços nos satélites SAR e nos algoritmos de processamento associados nas últimas décadas, a interferometria SAR emergiu como um método rotineiro para monitorar deformações do solo em larga escala e interpretar fenômenos geofísicos. Para uma análise eficaz das informações presentes na fase interferométrica, é fundamental reduzir o ruído de fase.

Métodos de filtragem desempenham um papel essencial no processamento de interferogramas, especialmente antes do desembrulhamento de fase. Esta técnica é utilizada para resolver a ambiguidade do módulo  $2\pi$  (ZEYADA et al., 2022) que surge devido à natureza periódica das medidas de fase obtidas pelo radar, permitindo reconstruir a continuidade da fase ao longo do espaço, eliminando os saltos de fase e fornecendo uma representação mais precisa e útil das informações contidas no interferograma.

A Organização Indiana de Pesquisa Espacial (ISRO – Indian Space Research Organization) e o Laboratório de Propulsão a Jato (JPL – Jet Propulsion Laboratory) da Administração Nacional da Aeronáutica e Espaço (NASA) estão desenvolvendo a missão NASA-ISRO Synthetic Aperture Radar (NISAR – NASA-ISRO Synthetic Aperture Radar): um satélite de imagem de radar de abertura sintética de dupla frequência (bandas L e S). Ele fornecerá dados InSAR em larga escala para uma ampla gama de ciências e aplicações (JONES; P; RAO, 2021). O NISAR produzirá cerca de 80 TB por dia, a maior quantidade de dados já liberada por uma missão de ciências da Terra em que a NASA esteve envolvida (SINGH, 2023).

Os dados do NISAR ajudarão as comunidades a se prepararem para perigos naturais e causados pelo homem, além de facilitar a resposta e recuperação após esses eventos. Por exemplo, o satélite será capaz de detectar mudanças sutis e lentas na forma e elevação da superfície terrestre, as quais podem preceder perigos naturais como terremotos, deslizamentos de terra e atividades vulcânicas. Após tais eventos, os dados do NISAR poderão ser usados para avaliar os danos no solo, fornecendo informações acionáveis para agências de resposta a desastres em um curto período de tempo (SINGH, 2023).

Diante da necessidade de técnicas de filtragem rápidas e eficientes para análise de grandes volumes de dados, propomos três novos filtros: o Extensão do Filtro Lee InSAR Refinado (LInSARRFE), fundamentado no modelo físico descrito em (3.20), e os filtros Filtro Normal Truncada (TcNfilter) e Filtro Cauchy Truncada (TcCfilter), baseados nos modelos empíricos Normal Truncada ( $\mathcal{NT}$ ) e Cauchy Truncada ( $\mathcal{CT}$ ), respectivamente. Esses filtros representam uma solução eficaz para o processamento e análise dos dados do NISAR, sendo particularmente úteis em aplicações que exigem agilidade na análise.

Para viabilizar a aplicação dessas metodologias, os métodos propostos foram implementados na plataforma R por meio dos pacotes `FDDPhase` e `Filters`. O primeiro é voltado para a análise estatística de dados de fase com modelos univariados físicos (circulares) e empíricos (lineares), enquanto o segundo disponibiliza os métodos de filtragem desenvolvidos a partir dos modelos presentes em `FDDPhase`.

Nesta pesquisa, avaliamos a eficiência dos métodos propostos, utilizamos três imagens: duas obtida pelo radar SAR do Sentinel-1, banda C, e uma gerada por simulação. Primeiramente, analisamos os dados simulados (VITALE et al., 2022). Em seguida, analisamos os dados da erupção do vulcão La Cumbre, localizado na Ilha Fernandina, arquipélago de Galápagos, Equador, ocorrido em 12 de janeiro de 2020 (PECHNIKOV, 2020). Por fim, estudamos a região que abrange a área afetada por um terremoto ocorrido na fronteira Irã-Iraque, com a cidade curda iraquiana de Halabja, em 12 de novembro de 2017 (PECHNIKOV, 2017).

## 1.1 OBJETIVOS

O objetivo geral desta pesquisa é aprimorar o conhecimento em análise estatística de dados circulares, com foco na avaliação da qualidade das imagens SAR/InSAR e no desenvolvimento de modelos e ferramentas para análise de dados de fase. O estudo busca apresentar novos métodos de filtragem de ruído de fase, desenvolvendo algoritmos baseados em modelos físicos e empíricos univariados para a diferença de fase. O intuito é expandir, analisar e comparar a eficácia desses métodos na redução de ruído, estabelecendo comparações com os filtros já existentes na literatura. Como objetivos específicos, temos:

1. Revisar a literatura sobre modelos estatísticos para análise de dados circulares, além de apresentar distribuições derivadas tanto de fundamentos físicos quanto empíricos, destacando suas relações e apresentando estimadores para seus parâmetros;
2. Avançar o estado da arte na avaliação da qualidade das imagens SAR/InSAR, propondo novos métodos de filtragem para a redução do ruído de fase, com foco no desenvolvimento de novos métodos e funcionalidades para a análise estatística de dados de fase.
3. Desenvolver, avaliar e disponibilizar, ao ambiente R, os pacotes `FDDPhase` e `Filters`. O primeiro destina-se à análise estatística para dados de fase, alinhando-se a modelos físicos e empíricos univariados. O segundo abrange métodos de filtragem aplicáveis a dados de fase;

4. Estender os pacotes `circular` e `CircStats` que tratam de modelos probabilísticos para análise de dados circulares, já existentes na literatura.

Por meio desses objetivos, esperamos contribuir para o desenvolvimento de técnicas precisas e eficientes em análise de dados circulares e na filtragem do ruído de fase.

## 1.2 PLATAFORMAS COMPUTACIONAIS

Para o desenvolvimento deste trabalho, utilizamos a linguagem R, idealizada para realizar análise de dados por meio de um sistema de computação estatística e gráfica (ALCOFORADO, 2021). Por outro lado, para o processo de desembrulhamento de fase, utilizamos o Python (MIGUEL et al., 2002), uma linguagem de programação amplamente empregada em uma variedade de domínios, incluindo aplicações da Web, desenvolvimento de software, ciência de dados e aprendizado de máquina.

Como recurso computacional para este trabalho utilizamos um computador com a seguinte configuração: DESKTOP-B11OC5E, sistema operacional Windows 10 Pro 64 bits, idioma português, modelo DELL Inspiron 3480, processador Intel(R) Core i5-8265U CPU @ 1.60GHz, 1.8GHz com 4096MB de memória RAM.

O código que implementa a proposta apresentada nesta pesquisa está disponível publicamente no GitHub, no link: <<https://github.com/araujoUFPE>>. Este repositório permite que outros pesquisadores possam reproduzir os experimentos e adaptar a metodologia para diferentes contextos.

## 1.3 TRABALHOS PUBLICADOS

Alguns dos resultados obtidos nesta pesquisa foram apresentados no 3rd e 4rd China International SAR Symposium (CISS), realizados em Xangai, na China em novembro de 2022 e dezembro de 2023, respectivamente. O CISS é uma conferência internacional dedicada às tecnologias e aplicações SAR. Posteriormente, esses resultados foram publicados pelo Institute of Electrical and Electronic Engineers (IEEE).

- FDDPhase: A Package For Statistical Analysis of the Phase from Polarimetric Data. DOI (Digital Object Identifier) [10.1109/CISS57580.2022.9971358](https://doi.org/10.1109/CISS57580.2022.9971358) e localizador URL (Uniform Resource Locator) <<https://ieeexplore.ieee.org/document/9971358>>;

- Fast Filtering of Interferometric Phase Noise in InSAR Data.  
DOI (Digital Object Identifier) [10.1109/CISS60136.2023.10380015](https://doi.org/10.1109/CISS60136.2023.10380015) e localizador URL (Uniform Resource Locator) <<https://ieeexplore.ieee.org/document/10380015>>.

## 1.4 ORGANIZAÇÃO DO TRABALHO

Estruturalmente, esta pesquisa está organizada em cinco capítulos.

- Capítulo 1: Apresentamos um breve histórico da evolução dos sistemas SAR, destacando suas principais características. Abordamos o impacto do ruído *speckle* nas imagens SAR, os métodos utilizados para sua redução e a relevância da modelagem estatística na análise de imagens SAR/InSAR;
- Capítulo 2: Exibimos os aspectos gerais e algumas características dos sistemas SAR, dos radares SAR polarimétricos e suas polarizações, conceitos relacionados ao ruído *speckle*, a interferometria SAR e a modelagem de dados PolSAR multilook;
- Capítulo 3: Abordamos conceitos e definições de alguns modelos estatísticos, tanto físicos, quanto empíricos que se caracterizam por analisar dados circulares, modelos probabilísticos para diferença entre fases, função de log-verossimilhança para os modelos físicos, relações entre modelos estatísticos circulares, métodos de filtragem para dados InSAR e divergência estatística;
- Capítulo 4: Apresentamos a descrição, análise e discussões dos resultados.
- Capítulo 5: Apresentamos as contribuições dos métodos propostos, destacando seus avanços e impactos. Além disso, destacamos as principais conclusões da pesquisa e propomos direções promissoras para investigações futuras.

## 2 BASE TEÓRICA E OPERACIONAL DE SISTEMAS SAR

Neste capítulo, apresentamos definições, características e propriedades relevantes dos sistemas e imagens SAR, incluindo aspectos gerais dos radares polarimétricos e interferométricos. Também discutimos o processo de modelagem para dados PolSAR multilook, além de outras informações importantes para compreender o objeto de estudo desta pesquisa.

### 2.1 SISTEMAS SAR/POLSAR E SUAS CARACTERÍSTICAS

Segundo (BORBA, 2020), nos sistemas SAR, os ecos das micro-ondas retornadas são capturados e representados como imagens. Essas imagens são geradas ao longo da direção azimutal, onde o radar emite um impulso em direção ao alvo e, em seguida, recebe e armazena o sinal de retorno, que varia conforme o espalhamento do sinal pelo alvo. A capacidade dos sistemas SAR de operar em diferentes condições climáticas, aliada à sua eficiência em gerar imagens com alta resolução espacial, torna-os uma ferramenta valiosa para o sensoriamento remoto.

Assim como o SAR, a polarimetria é uma técnica amplamente empregada no mapeamento de recursos terrestres. Na polarimetria os dados são representados por matrizes de espalhamento, as quais estabelecem relações entre a energia incidente e a energia retroespalhada as várias combinações de polarizações. O PolSAR, descreve as relações de fase entre os pulsos transmitidos e recebidos. De forma resumida, a polarimetria concentra-se em estudar:

- A orientação da onda do radar;
- O espalhamento da onda do radar;
- A informação de fase entre os canais.

Os dados provenientes de imagens SAR são gravados como uma faixa de posições para cada tempo avançado na direção azimutal ou rota (BORBA, 2020), (LEE; POTTIER, 2009). Essas imagens são adquiridas por radares que operam na faixa de micro-ondas e, por causa disso, apresentam informações adicionais em comparação às imagens óticas. Por exemplo, os radares SAR têm a capacidade de penetrar nuvens e de operar durante a noite, além de detectar variações no relevo, movimentos e deformações do solo. As pesquisas utilizando imagens SAR continuam em crescente avanço em sensoriamento remoto (LEE; POTTIER, 2009).

Nos sistemas SAR, as informações no retroespalhamento podem ser representadas por imagens. Dados SAR/PolSAR são obtidas por meio do envio de pulsos em direção ao alvo em cada ponto ao longo da direção azimutal. Durante a aquisição das imagens, as informações são armazenadas em uma matriz de dados associadas as direções de range e azimute.

## 2.2 RADARES SAR POLARIMÉTRICOS E POLARIZAÇÃO

O princípio básico de um radar de imagem consiste em emitir um sinal eletromagnético em direção a uma superfície e registrar tanto a quantidade de sinal refletido, conhecida como backscatter, quanto o atraso no tempo de retorno. O processador SAR armazena esses sinais, registrando suas amplitudes e fases, que são posteriormente processadas para gerar imagens. A construção das imagens SAR baseia-se na intensidade e no tempo de retorno do sinal, os quais estão diretamente relacionados à rugosidade da superfície, às suas propriedades elétricas e à distância entre o radar e a área observada.

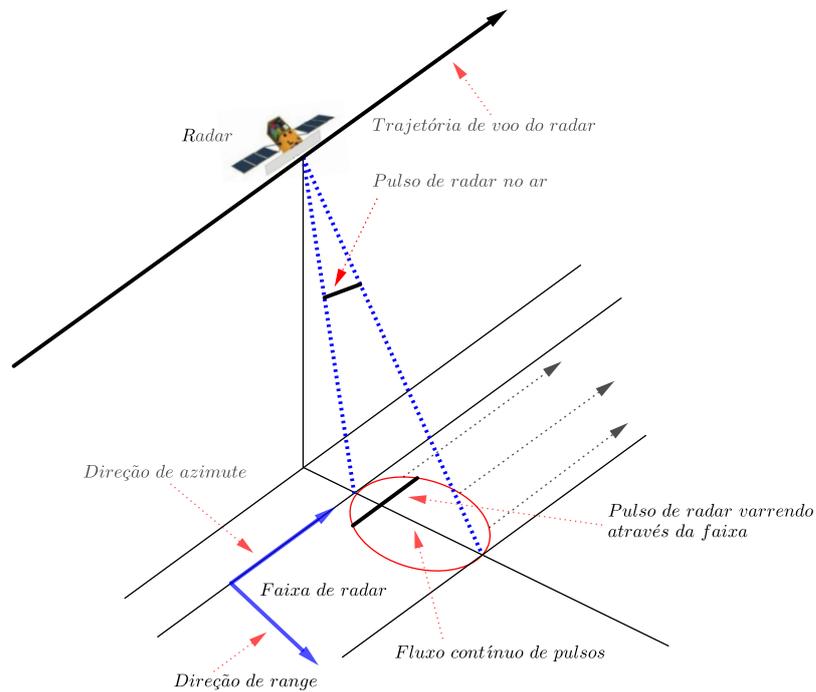
Para garantir imagens com alta resolução, os sistemas SAR (que podem ser satélites ou aerotransportados) empregam diferentes técnicas nas direções de range e azimute. Na direção de range, a alta resolução é alcançada por meio de pulsos com grande largura de banda, definidos pela variação da frequência transmitida. Já na direção azimutal, a técnica de abertura sintética permite obter alta resolução ao longo do deslocamento do sensor. A Figura 1 apresenta a configuração geométrica básica de um sistema de imageamento por Radar de Abertura Sintética.

Na configuração de um radar, conforme ilustrado na Figura 1, a antena emite pulsos que iluminam uma área específica no solo determinada pelo comprimento de onda e dimensão da antena. Os pulsos são enviados e recebidos de forma contínua, o que permite que cada ponto no solo seja amostrado várias vezes ao longo do tempo.

Uma antena de radar pode ser projetada para enviar e receber ondas eletromagnéticas com uma polarização previamente definida. Sistemas SAR podem reunir informações detalhadas sobre as propriedades polarimétricas da superfície observada, que podem revelar a estrutura, orientação e condições ambientais dos elementos da superfície. Por exemplo, estruturas orientadas linearmente, como edifícios ou ondulações na areia, tendem a refletir e preservar a coerência, ou seja, mesma direção linear do sinal polarimétrico.

Uma onda transmitida nas polarizações H ou V resulta em uma onda retroespalhada com polarização variável. O sistema SAR polarimétrico é caracterizado pelas diferentes combinações

Figura 1 – Geometria básica de um sistema de imageamento por Radar de Abertura Sintética



Fonte: O autor (2025).

de polarizações transmitidas e recebidas, ou segundo Martínez e Pottier (2021), a polarimetria refere-se especificamente à natureza vetorial das ondas eletromagnéticas, enquanto a polarimetria de radar é a ciência de adquirir, processar e analisar o estado de polarização da onda eletromagnética em aplicações de radar.

Teoricamente, dados SAR polarimétricos permitem extrair o máximo de informação da resposta em retroespalhamento do alvo iluminado, quando comparado com aquisições em polarização simples, dual ou mesmo multipolarizada. O PolSAR registra as mudanças de estado de polarização do sinal retroespalhado pelas características elétricas e geométricas do alvo, de tal forma, que estas propriedades podem ser derivadas através da informação de amplitude de cada polarização (componentes HH, VV, HV, VH) e da fase entre os componentes (PARADELLA et al., 2015).

A direção de polarização é definida pela orientação do campo elétrico da onda, que é sempre perpendicular ao seu campo magnético. As medições de radar podem ser polarizadas como:

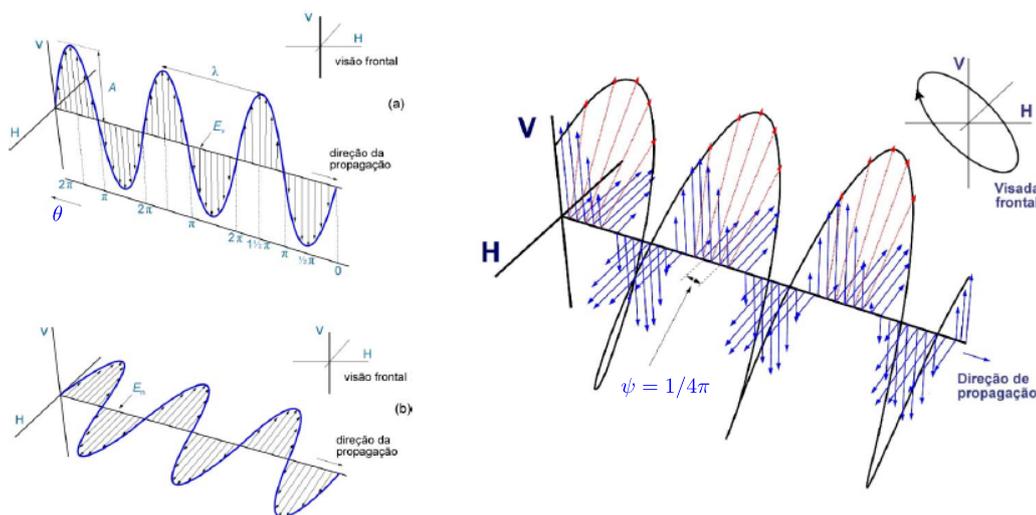
- HH: Transmitido horizontalmente e recebido horizontalmente;
- HV: Transmitido horizontalmente e recebido verticalmente;
- VH: Transmitido verticalmente e recebido horizontalmente;

- VV: Transmitido verticalmente e recebido verticalmente.

A radiação eletromagnética que se propaga possui três campos vetoriais associados a ela, são eles: *i*) campo elétrico, *ii*) campo magnético e *iii*) vetor de propagação da onda. Para ondas planas e uniformes estes três vetores são mutualmente ortogonais. Em geral, podemos definir a polarização de uma onda eletromagnética pela forma geométrica que o vetor campo elétrico descreve no espaço, ou seja, elipse, círculo ou linear, as quais possuem polarizações, elíptica, circular ou linear, nesta ordem.

As polarizações HH e VV estão mais relacionadas com a rugosidade da superfície (espalhamento superficial), enquanto, as polarizações, HV e VH, estão mais relacionadas com o interior dos objetos (espalhamento volumétrico). Os sistemas de radar podem ter combinações de polarizações distintas, ou seja, polarização simples (HH ou VV); polarização dupla (HH e HV, VV e VH ou HH e VV) e completamente polarimétrica (HH, VV, HV e VH) e ainda, um dos componentes mais importante de um radar polarimétrico, o ângulo associado ao retorno complexo, ou seja, a fase relativa entre estas polarizações.

Figura 2 – Direções de propagação do campo elétrico: (a) polarização vertical (V) e (b) polarização horizontal (H)



Fonte: (PARADELLA et al., 2005)

Na Figura 2, podemos identificar: *(i)* Atributos da onda: comprimento ( $\lambda$ ), amplitude ( $A$ ), fase ( $\theta$ ), campo elétrico vertical ( $E_V$ ) e campo elétrico horizontal ( $E_H$ ); *(ii)* Polarização da onda: para uma diferença de fase entre dois componentes polarizados,  $\psi = 1/4\pi$ , resulta em uma onda com polarização elíptica (PARADELLA et al., 2005).

As imagens adquiridas por radares polarimétricos são formadas por micro-ondas de retorno, combinando diferentes polarizações. Esse processo de polarização permite obter mais informações sobre a região (alvo) em comparação com as informações obtidas por imagens SAR convencionais. Dessa forma, dados polarimétricos têm o propósito de aprimorar a compreensão e o entendimento do espalhamento das micro-ondas pelos alvos, considerando as diferentes polarizações. Por exemplo, distinguir tipos de cobertura no solo, como áreas urbanas, florestas, campos agrícolas. Além disso, em aplicações marítimas, diferenciar entre água calma e ondas agitadas ou até mesmo para identificar a presença de óleo na superfície do mar.

### 2.3 RUÍDO SPECKLE

Uma das características inerentes às imagens SAR é a presença de ruído speckle, o qual resulta em variações nas intensidades dos pixels da imagem. Esse ruído produz um padrão granular que dificulta a interpretação e análise das imagens em relação à tipologia da área imageada. O aparecimento do speckle em imagens SAR deve-se à interferência coerente de ondas refletidas por muitos espalhadores elementares. Deste modo, o conhecimento preciso das propriedades estatísticas do speckle é importante para o processamento e entendimento de uma imagem SAR (LEE; POTTIER, 2009). Essas propriedades podem ser utilizadas para discriminar os tipos de terrenos utilizados e para desenvolver filtros especializados para a redução do ruído speckle (FRERY et al., 1997).

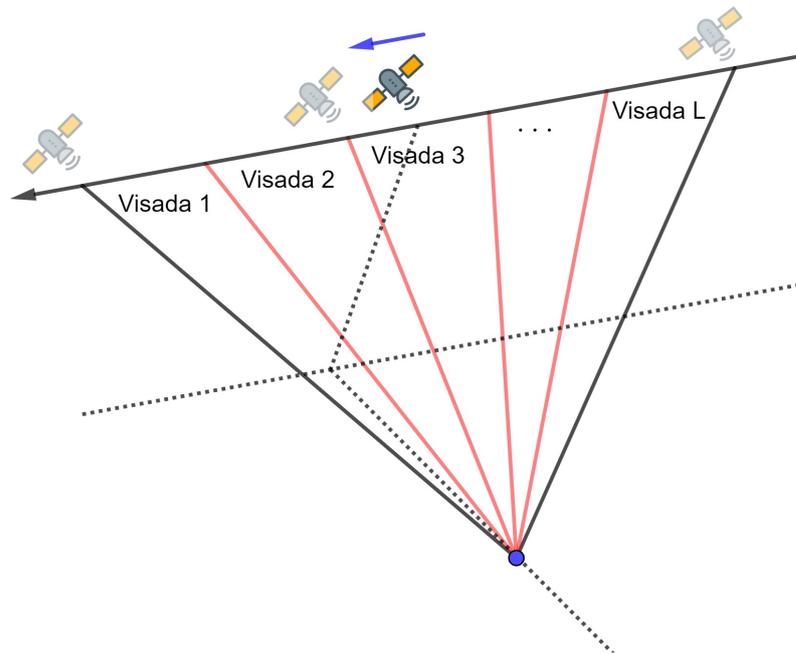
Em uma imagem SAR, um pixel pode ser descrito por uma matriz aleatória  $Z^{(1)}$  (single-look), na qual, cada elemento dessa matriz representa a intensidade do sinal de retroespalhamento do terreno em uma determinada polarização. A matriz de retorno  $Z^{(1)}$  é, geralmente, modelada como  $Z^{(1)} = X \cdot Y$ , onde  $X$  é uma variável escalar positiva que modela o retroespalhamento (representa a intensidade do sinal retroespalhado) com  $\mathbb{E}[X] = 1$  e  $Y$  é uma variável complexa, que modela os dados do ruído speckle.

Nos sistemas SAR, os dados das imagens são obtidos com a soma coerente (vetorial) do sinal retornado por cada um dos elementos difusores que compõem uma célula de resolução da cena analisada. Como os elementos difusores estão distribuídos de forma aleatória, a soma vetorial dos sinais de retorno sofre interferências tanto construtivas quanto destrutivas, resultando em imagens com aspecto granuloso (HORTA, 2009).

Uma abordagem comum para reduzir o ruído speckle é o processamento multilook. De acordo com Lee e Pottier (2009), nessa técnica, cada segmento é processado independente-

mente para formar uma imagem SAR de intensidade ou amplitude, e as  $L$  imagens são somadas para formar uma imagem SAR de  $L$  looks. Esse processo melhora a qualidade das imagens SAR ao calcular a média de várias amostras ou visadas. Ao reduzir o ruído, o multilook suaviza a imagem e gera pixels mais uniformes e quadrados na imagem de saída. Como podemos observar na Figura 3, o multilook consiste em dividir o comprimento da abertura sintética em visadas, ou seja, em  $L$  segmentos ou looks.

Figura 3 – Divisão da abertura sintética em segmentos ou visadas



Fonte: O autor (2025).

Esta técnica aumenta a relação sinal-ruído da imagem final, diminuindo o efeito do ruído speckle. No entanto, o processamento multilook provoca uma degradação na resolução em azimute, uma vez que as imagens de cada look possuem uma largura de banda menor que a largura total, ou seja, a imagem tende a tornar-se mais suavizada.

Para imagens PolSAR multilook e admitindo o modelo multiplicativo, podemos representar os dados do ruído speckle por uma matriz de covariância complexa com  $L$  looks (seção 2.5) e ordem  $q$ , no qual  $q$  é o número de polarizações. Outras informações sobre, formação, propriedades, descrição estatística e filtragem do ruído speckle podem ser consultadas em (LEE; POTTIER, 2009) e (MARTÍNEZ; POTTIER, 2021).

## 2.4 INTERFEROMETRIA SAR

Os avanços nas pesquisas em sensoriamento remoto têm sido notáveis, impulsionados pela crescente disponibilidade de dados orbitais SAR adquiridos sistematicamente. Essa evolução tecnológica permite medir com precisão a distância entre o sensor e o alvo, utilizando o intervalo de tempo entre a emissão das ondas eletromagnéticas pela antena e a recepção do sinal retroespalhado da área imageada. Segundo Gama et al. (2013), as variações na distância medida em momentos diferentes permitem a detecção de deslocamento da Terra, sendo o princípio da técnica interferométrica SAR.

A técnica InSAR combina duas ou mais imagens SAR sobre a mesma região para revelar a topografia ou o movimento da superfície. A interferometria analisa a fase dos sinais de radar recebidos por duas antenas localizadas em posições diferentes ou por uma mesma antena em época e posição diferente. Se as imagens SAR combinadas se originarem de posições diferentes, a topografia da superfície pode ser mapeada. Por outro lado, se imagens da mesma posição, mas obtidas em épocas diferentes, forem combinadas, a diferença entre elas mostrará o movimento ou a deformação da superfície no tempo entre as duas imagens. A imagem produzida é conhecida como interferograma.

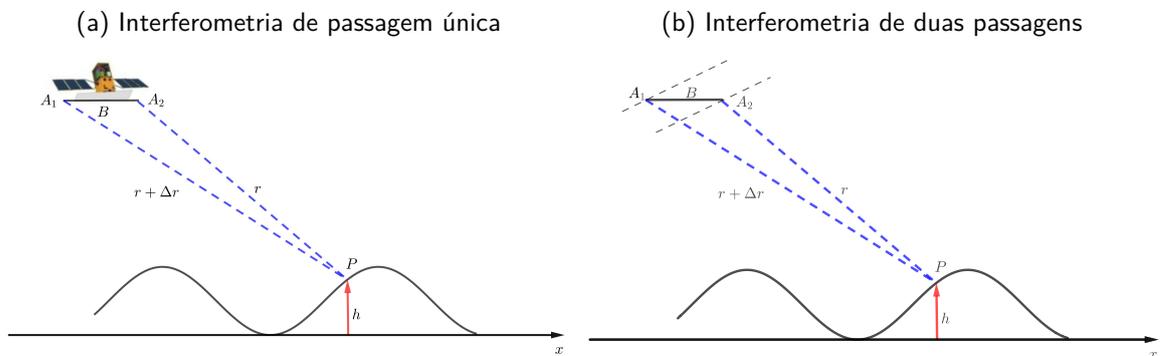
A interferometria SAR é a técnica que utiliza um par de imagens SAR em formato complexo (amplitude e fase) para gerar uma terceira imagem complexa, denominada imagem interferométrica ou interferograma. Nessa imagem, cada pixel contém informação sobre a fase interferométrica, ou seja, sobre a diferença de fase entre os pixels correspondentes nas duas imagens originais. A fase de cada pixel está relacionada à elevação da superfície terrestre, o que permite que sejam gerados modelos numéricos do terreno (MURA, 2000).

Ao contrário do SAR, que utiliza tanto a amplitude de uma onda (intensidade do sinal) quanto a fase do sinal retornado, para a interferometria a medida mais importante é a diferença de fase ou fase interferométrica, ou seja, a interferometria SAR se concentra na análise da fase dos sinais de radar recebidos, e a partir desses sinais, podemos gerar imagens de amplitude e fase em formato complexo, adquiridas sobre a mesma área, a partir de dois ângulos de aquisição diferentes, com objetivo de gerar um interferograma. O interferograma representa a fase interferométrica e através desta é possível gerar o modelo de elevação do terreno (MURA, 2005). A geometria de aquisição de uma imagem SAR pode ser caracterizada por:

- i*) Interferometria de passagem única (single pass interferometry): utiliza-se duas antenas,  $A_1$  e  $A_2$ , montadas na mesma plataforma numa direção perpendicular à direção de imageamento (azimute) e separadas por uma distância  $B$ , chamada de linha base.
- ii*) Interferometria de duas passagens (two pass interferometry): utiliza-se uma única antena e realiza duas passagens paralelas em diferentes tempos sobre a mesma área.

Os casos *i*) e *ii*) são ilustrados na Figura 4, onde a diferença nas distâncias entre o alvo ( $P$ , ponto genérico) e as antenas  $A_1$  e  $A_2$  é, respectivamente,  $A_1$  e  $A_2$ , sendo  $\Delta r$  a variação de distância e  $h$  a elevação do terreno.

Figura 4 – Interferometria de passagem única e dupla



Fonte: O autor (2025).

### 2.4.1 Processo de aquisição de imagens SAR

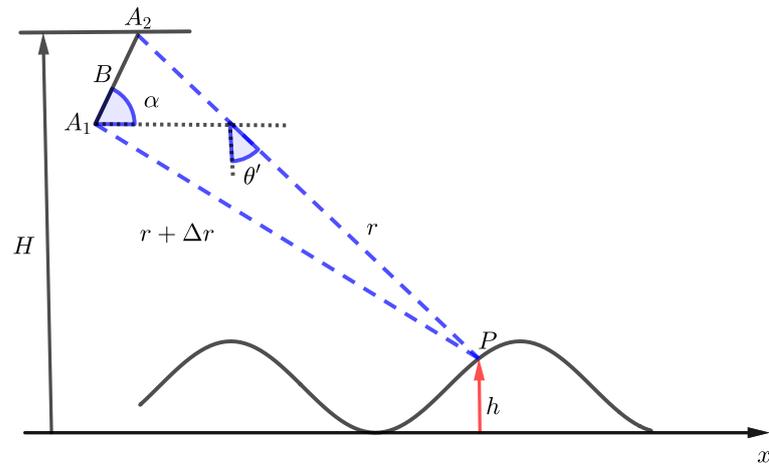
A Figura 5 apresenta uma ilustração da geometria envolvida no processo de aquisição de imagens SAR. Nessa geometria, uma antena emite pulsos de micro-ondas que são direcionados para o alvo de interesse no solo. O sinal refletido pelo alvo é então recebido pela antena e registrado como o sinal de retroespalhamento. A distância entre a antena e o alvo é determinada pelo tempo que o sinal leva para percorrer o caminho de ida e volta.

A relação entre a diferença de fase  $\psi$ , os dois sinais de onda de comprimento  $\lambda$  e a diferença das distâncias percorridas do alvo às antenas  $A_1$  e  $A_2$  é expressa pela seguinte equação:

$$\psi = \frac{4\pi\Delta r}{\lambda}. \quad (2.1)$$

A interferometria SAR utiliza informações de  $\psi$ , obtidas por meio da fase interferométrica entre duas imagens SAR para produzir informações de  $\Delta r$ , obtendo assim, o valor de  $h$ , ou

Figura 5 – Geometria básica de um sistema InSAR



Fonte: O autor (2025).

seja, o valor de elevação do ponto genérico  $P$ . De acordo com o triângulo de vértices  $A_1$ ,  $A_2$  e  $P$ , da Figura 5, conhecendo o ângulo de incidência  $\theta'$  e a distância  $r$  de um ponto qualquer  $P$ , chega-se ao valor de elevação do ponto  $P$ , dado por:

$$h(r, \theta') = H - r \cos \theta' \quad (2.2)$$

onde,  $H$  representa a altura da plataforma.

#### 2.4.2 Diferença de fase ou fase interferométrica

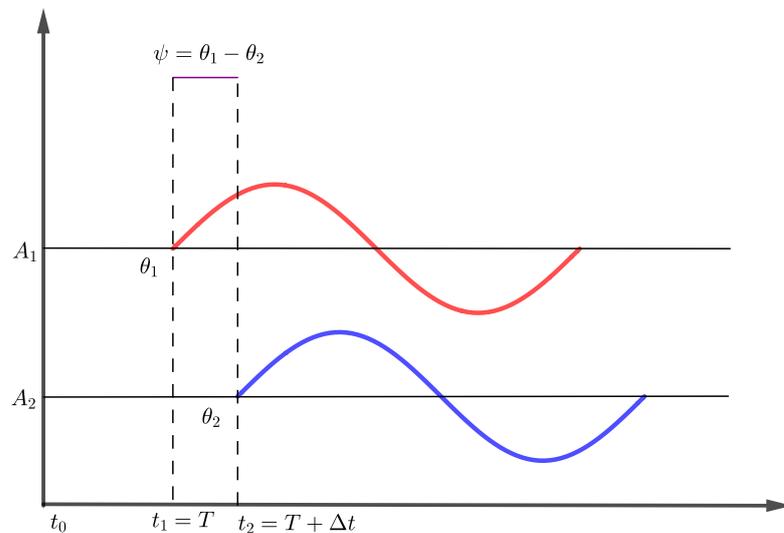
A partir de dados de natureza complexa, é possível gerar uma nova imagem na qual cada pixel é formado pela diferença de fase entre os pixels correlatos ou homólogos nas duas imagens originais. A diferença de fase representa o desvio, no tempo ou no espaço, entre duas ondas. As ondas estão em fase quando suas origens de fase estiverem perfeitamente alinhadas. Por exemplo, duas ondas estão em fase se começarem em  $-\pi$  e terminarem em  $\pi$  (ou, começar em 0 e terminar em  $2\pi$ ) ao mesmo tempo, ou seja, situação em que as cristas de duas ondas se alinham, de modo que as duas ondas se somam. Quando isto não ocorre, dizemos que as ondas estão fora de fase.

A diferença de fase entre pixels de mesma coordenada nas imagens está relacionada à distância entre as duas trajetórias de aquisição, linha base e a altitude correspondente no solo. Combinando interferometria e polarimetria é possível maximizar a coerência complexa entre as duas imagens, e por consequência a melhoria da fase interferométrica, visando a geração de modelos de elevação mais precisos (MURA, 2005).

A fase interferométrica é determinada pela diferença de fase entre dois sinais refletidos, que são combinados de maneira coerente. Essa diferença de fase é limitada ao intervalo  $(-\pi, \pi)$  devido à natureza do elemento de resolução complexa, que restringe os valores ao módulo  $2\pi$ . Quando a diferença de fase ultrapassa esse intervalo, ela é embrulhada ou enrolada de volta dentro dos limites definidos.

A diferença total de intervalos entre dois pontos de observação, que representa a fase, em geral, podem ser muitos múltiplos do comprimento de onda do radar ou, expressos em termos de fase, múltiplos de  $2\pi$  (GAMA et al., 2013). Isto ocorre, devido à natureza periódica da fase, por este motivo, a fase do interferograma é conhecida como fase embrulhada (wrapped phase). De acordo com Sica et al. (2021), o desembrulhamento de fase é o processo de recuperar a fase contínua de um sinal interferométrico a partir de medições de fase discretas, que estão sujeitas a ambiguidades devido à natureza cíclica da fase. Na Figura 6, podemos ver como é obtido a diferença de fase para duas imagens SAR.

Figura 6 – Diferença de fase ou fase interferométrica



Fonte: O autor (2025).

Na interferometria, a obtenção da fase interferométrica começa com o registro das imagens complexas, garantindo que os pixels em posições correspondentes nas duas imagens representem o mesmo elemento de resolução no solo. Após esse registro, o interferograma é gerado pelo produto dos pixels da primeira imagem com os valores conjugados correspondentes da segunda imagem. Para ilustrar o processo, considere os seguintes valores complexos para os pixels correspondentes das imagens 1 e 2 (Figura 7):

- $a_{i,j} = \rho_1 \exp(i\theta_1)$ , que representa o valor do pixel na imagem 1, onde  $\rho_1$  é a amplitude e  $\theta_1$  é a fase do sinal naquele pixel.
- $b_{i,j} = \rho_2 \exp(i\theta_2)$ , que representa o valor do pixel na imagem 2, onde  $\rho_2$  é a amplitude e  $\theta_2$  é a fase do sinal naquele pixel.

A operação para gerar um interferograma é dada pela multiplicação do pixel  $a_{i,j}$  da primeira imagem pelo conjugado complexo do pixel  $b_{i,j}$  da segunda imagem. O conjugado complexo de  $b_{i,j}$  é dado por  $b_{i,j}^* = \rho_2 \exp(-i\theta_2)$ , ou seja, ele possui a mesma parte real de  $b_{i,j}$ , no entanto, com parte imaginária simétrica. Por definição, o produto entre  $a_{i,j}$  e  $b_{i,j}^*$  resulta em:

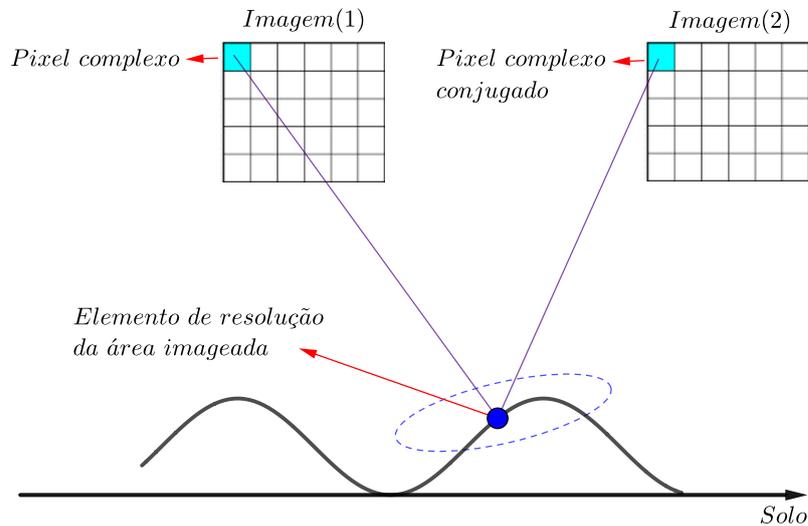
$$c_{i,j} = \rho_1 \exp(i\theta_1) \cdot \rho_2 \exp(-i\theta_2).$$

Esse produto pode ser simplificado como:

$$c_{i,j} = \rho_1 \rho_2 \exp(i(\theta_1 - \theta_2)) = \rho_1 \rho_2 \exp(i\psi),$$

onde,  $\psi = \theta_1 - \theta_2$  é a diferença de fase e  $*$  denota o transposto conjugado.

Figura 7 – Representação de duas imagens complexas co-registradas



Fonte: O autor (2025).

## 2.5 MODELAGEM PARA DADOS POLSAR MULTILOOK

Sistemas polarimétricos medem as propriedades de um alvo, detectando a mudança no estado de polarização que o alvo provoca na onda incidente, isto é, os pulsos de micro-ondas transmitidos por sistemas PolSAR medem as componentes do sinal recebido, onde para

cada pixel, a medida resultante é uma matriz de espalhamento com coeficientes complexos. Adotando,  $\underline{\mathbf{E}}^i$  como a onda incidente polarizada e  $\underline{\mathbf{E}}^s$  como a onda retroespalhada recebida, podemos escrever (MARTÍNEZ; POTTIER, 2021):

$$\underline{\mathbf{E}}^i = \begin{bmatrix} E_{\text{H}}^i \\ E_{\text{V}}^i \end{bmatrix}, \quad \underline{\mathbf{E}}^s = \begin{bmatrix} E_{\text{H}}^s \\ E_{\text{V}}^s \end{bmatrix}, \quad (2.3)$$

em que H e V representam os estados de polarização horizontal e vertical, respectivamente, enquanto os sobrescritos  $i, s$  indicam a onda incidida e recebida, nesta ordem.

É possível estabelecer a relação entre a onda espalhada  $\underline{\mathbf{E}}^s$  e a onda incidente  $\underline{\mathbf{E}}^i$  por meio de uma matriz, expressa por:

$$\underline{\mathbf{E}}^s = \frac{e^{-jkd}}{d} \mathbf{S} \underline{\mathbf{E}}^i. \quad (2.4)$$

Em (2.4),  $d$  é a distância entre o alvo e a antena receptora e  $k$  é o número de onda. O termo  $e^{-jkd}/d$  leva em consideração os efeitos de propagação da onda, tanto em amplitude, quanto em fase. Por outro lado, o fator de fase representa o atraso do percurso da onda do espalhador até a antena, o qual, é medido em radianos e descreve a diferença de fase entre a onda e um sinal de referência. Assim, o processo de dispersão que ocorre no alvo de interesse expresso em (2.4) pode ser reescrita como:

$$\begin{bmatrix} E_{\text{H}}^s \\ E_{\text{V}}^s \end{bmatrix} = \frac{e^{-jkd}}{d} \mathbf{S} \begin{bmatrix} E_{\text{H}}^i \\ E_{\text{V}}^i \end{bmatrix}. \quad (2.5)$$

Um radar polarimétrico mede o scattering (espalhamento) da matriz  $\mathbf{S}$  de um meio, em um determinado ângulo de incidência (LEE et al., 1994). Neste caso, com radares polarimétricos podemos definir os retornos em cada polarização da onda transmitida e recebida, formando com os alvos, uma matriz de espalhamento complexo, dada por:

$$\mathbf{S} = \begin{bmatrix} S_{\text{HH}} & S_{\text{HV}} \\ S_{\text{VH}} & S_{\text{VV}} \end{bmatrix}, \quad (2.6)$$

em que  $S_i, S_j$  são quaisquer dois componentes de dispersão da matriz  $\mathbf{S}$  ou dois retornos de radares SAR polarimétricos ou interferométricos, com  $\{i, j\} \in \{\text{HH}, \text{HV}, \text{VH}, \text{VV}\}$ .

Por definição, a matriz  $\mathbf{S}$  depende do sistema de coordenadas. Quando as antenas transmissora e receptora estão no mesmo local, consideramos que as medições são monostáticas e utilizamos o sistema de Alinhamento de Espalhamento Traseiro (BSA – Back Scattering

Alignment). Ou seja, os canais cross-pol da matriz de espalhamento  $\mathbf{S}$  são iguais, então, a formalização do processo de espalhamento dado em (2.5), no caso monostático sob a convenção BSA, pode ser reduzidos a:

$$\begin{bmatrix} E_{\text{H}}^s \\ E_{\text{V}}^s \end{bmatrix} = \frac{e^{-jkd}}{d} \begin{bmatrix} S_{\text{HH}} & S_{\text{HV}} \\ S_{\text{HV}} & S_{\text{VV}} \end{bmatrix} \begin{bmatrix} E_{\text{H}}^i \\ E_{\text{V}}^i \end{bmatrix}. \quad (2.7)$$

Um parâmetro central, ao se considerar o processo de espalhamento que ocorre em um determinado espalhador, consiste na potência espalhada. Para sistemas de polarização simples, a potência espalhada é determinada por meio da Seção Transversal do Radar (RCS – Radar Cross-Section) ou do coeficiente de espalhamento. No entanto, um radar polarimétrico deve ser considerado como um sistema multicanal. Portanto, para determinar a potência espalhada, é necessário considerar todos os canais de dados, ou seja, todos os elementos da matriz  $\mathbf{S}$ . A potência total espalhada, no caso de um sistema de radar polarimétrico é conhecida como Span, sendo definida no caso mais geral da seguinte forma:

$$\text{Span}(\mathbf{S}) = \text{tr}(\mathbf{S}\mathbf{S}^*) = |S_{\text{HH}}|^2 + |S_{\text{HV}}|^2 + |S_{\text{VH}}|^2 + |S_{\text{VV}}|^2, \quad (2.8)$$

em que  $\text{tr}(\cdot)$  é o traço da matriz.

No caso do retroespalhamento, devido ao teorema da reciprocidade vetorial, que afirma: se transmitirmos um estado de polarização  $P_A$  da posição  $A$ , então o componente polarizado na direção  $P_B$  na posição  $B$  é igual ao componente  $P_A$  da radiação dispersa quando iluminamos o mesmo objeto de  $B$  com polarização  $P_B$ . O Span se reduz a:

$$\text{Span}(\mathbf{S}) = \text{tr}(\mathbf{S}\mathbf{S}^*) = |S_{\text{HH}}|^2 + 2|S_{\text{HV}}|^2 + |S_{\text{VV}}|^2. \quad (2.9)$$

O Span é polarimetricamente invariável, ou seja, não depende da base de polarização empregada para descrever a polarização das ondas eletromagnéticas, esta, se caracterizando como a sua principal propriedade.

Com o objetivo de extrair informação física de um alvo, é possível transformar a matriz  $\mathbf{S}$  em um vetor de espalhamento, o qual, apresenta uma interpretação física mais clara, ou seja, nesse caso, a informação polarimétrica pode ser representada com um vetor complexo, denominado vetor de dispersão, dado por;

$$\mathbf{k} = \frac{1}{2} \text{tr}(\mathbf{S}\mathbf{\Psi}) \quad (2.10)$$

em que  $\mathbf{\Psi}$  é um conjunto completo de matrizes complexas  $2 \times 2$  que forma uma base para o espaço vetorial de matrizes Hermitianas.

Há diferentes bases, embora as mais utilizadas no contexto de polarimetria sejam a de Pauli e a Lexicográfica que geram, respectivamente, as matrizes de coerência e covariância. Seja  $\Psi_p$  o conjunto das matrizes de coerência polarimétrica da base de Pauli, no caso monostático pode-se definir:

$$\Psi_p = \left\{ \sqrt{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \sqrt{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \sqrt{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}, \quad (2.11)$$

em que (2.11) se reduz a um vetor de espalhamento, denominado vetor característico de Pauli 3-D, dado por:

$$\mathbf{k} = \frac{1}{\sqrt{2}} [S_{HH} + S_{VV} \quad S_{HH} - S_{VV} \quad 2S_{HV}]^T, \quad (2.12)$$

onde,  $T$  representa a transposta.

Seja  $\Psi_\ell$  o conjunto das matrizes de coerência polarimétrica da base Lexicográfica, o qual define-se como:

$$\Psi_\ell = \left\{ 2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, 2\sqrt{2} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, 2 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}, \quad (2.13)$$

em que representamos a matriz de espalhamento pelo vetor característico lexicográfico 3-D, da seguinte forma:

$$\Omega = [S_{HH} \quad \sqrt{2}S_{HV} \quad S_{VV}]^T. \quad (2.14)$$

Os termos  $2, \sqrt{2}$  e  $2\sqrt{2}$  em (2.11) e (2.13) são adicionados para manter a norma dos vetores de espalhamento iguais, independentemente da base escolhida, em outras palavras, preserva a potência total (2.9) do sinal, isto é:

$$\text{Span}(\mathbf{S}) = \text{tr}(\mathbf{S}\mathbf{S}^*) = |\mathbf{k}|^2 = |\Omega|^2 = |S_{HH}|^2 + 2|S_{HV}|^2 + |S_{VV}|^2. \quad (2.15)$$

De acordo com as formas vetoriais da matriz  $\mathbf{S}$ , os vetores característicos em (2.12) para a base de Pauli, e (2.14) para a base Lexicográfica, geram as matrizes de coerência  $\mathbf{T}_3$  e covariância  $\mathbf{C}_3$ , respectivamente, as quais, são definidas da seguinte forma:

$$\mathbf{T}_3 = \langle \mathbf{k}\mathbf{k}^* \rangle, \quad \mathbf{C}_3 = \langle \Omega\Omega^* \rangle, \quad (2.16)$$

em que  $\langle \cdot \rangle$  indica a média espacial ou temporal.

Para imagens SAR polarimétricas, cada entrada está relacionada aos elementos da matriz  $\mathbf{S}$ , ou seja, números complexos que representam os canais de polarização. Os retornos  $S_{HH}$  e  $S_{VV}$  são chamados de copolarizados, enquanto os retornos  $S_{VH}$  e  $S_{HV}$  são denotados como componentes de polarização cruzada. Teoricamente, devido a condição de reciprocidade<sup>1</sup> (LEE; POTTIER, 2009), os componentes  $S_{VH}$  e  $S_{HV}$ , coincidem. Deste modo, a matriz  $\mathbf{S}$ , pode ser reescrita da seguinte forma:

$$\mathbf{U} = \begin{bmatrix} S_{HH} \\ S_{VH} \\ S_{VV} \end{bmatrix}. \quad (2.17)$$

O vetor  $\mathbf{U}$  segue uma distribuição Gaussiana complexa multivariada de média nula. Uma vantagem da distribuição Gaussiana, se dá, ao fato, dela ser matematicamente tratável, permitindo assim a obtenção de resultados interessantes. Mas, estatisticamente, podemos indicar duas outras razões para o uso da distribuição Gaussiana: (i) distribuições amostrais de muitas estatísticas multivariadas são aproximadamente normais, devido ao efeito do Teorema do Limite Central; (ii) a distribuição Gaussiana serve como modelo aproximado em muitos problemas reais.

Dados PolSAR caracterizados na forma do vetor  $\mathbf{U}$  são denotados single look. No entanto, este tipo de dado, não leva em consideração o controle do ruído speckle sobre a imagem. Todavia, para termos o controle deste ruído, utilizamos o processamento multilook, o qual é realizado com a finalidade de possibilitar a visualização dos dados PolSAR. Apesar da natureza complexa dos dados, apenas a amplitude ou a intensidade, contém informações úteis sobre o alvo em análise (MARTÍNEZ; POTTIER, 2021).

O processamento SAR polarimétrico multilook consiste em calcular a média de várias matrizes de covariância independentes, provenientes de aquisições single look. A matriz de covariância amostral para  $L$  looks é definida por:

$$\mathbf{Z} = \frac{1}{L} \sum_{\ell=1}^L \mathbf{U}(\ell) \mathbf{U}^*(\ell) \quad (2.18)$$

<sup>1</sup> O Teorema da Reciprocidade estabelece que a resposta do sistema é simétrica sob a troca das posições do transmissor e do receptor. Ou seja, a matriz  $\mathbf{S}$  obedece à relação  $S_{VH} = S_{HV}$ .

em que  $\ell = 1, \dots, L$ , o superíndice  $*$  denota o transposto conjugado e,

$$\begin{aligned}
 \mathbf{U}(\ell)\mathbf{U}^*(\ell) &= \begin{bmatrix} S_{HH} \\ S_{VH} \\ S_{VV} \end{bmatrix} (\ell) \begin{bmatrix} S_{HH}^* & S_{VH}^* & S_{VV}^* \end{bmatrix} (\ell) \\
 &= \begin{bmatrix} S_{HH}S_{HH}^* & S_{HH}S_{VH}^* & S_{HH}S_{VV}^* \\ S_{VH}S_{HH}^* & S_{VH}S_{VH}^* & S_{VH}S_{VV}^* \\ S_{VV}S_{HH}^* & S_{VV}S_{VH}^* & S_{VV}S_{VV}^* \end{bmatrix} (\ell). \\
 &= \begin{bmatrix} I_{HH} & T_{HH,VH} & T_{HH,VV} \\ T_{HH,VH} & I_{VH} & T_{VH,VV} \\ T_{HH,VV} & T_{VH,VV} & I_{VV} \end{bmatrix} (\ell). \tag{2.19}
 \end{aligned}$$

Os elementos na diagonal principal de (2.19) são variáveis aleatórias positivas que modelam a intensidade  $I$  de cada canal. Enquanto, os elementos fora da diagonal principal são variáveis aleatórias complexas que descrevem a covariância  $T$  entre os canais. O formato resultante de  $\mathbf{Z}$  também é conhecido como Matriz de Covariância Amostral. Além disso, como os elementos fora da diagonal principal são conjugados complexos entre si, a matriz de covariância  $\Sigma$ , é definida por:

$$\Sigma = E[\mathbf{U}(\ell)\mathbf{U}^*(\ell)] = \begin{bmatrix} E[S_{HH}S_{HH}^*] & E[S_{HV}S_{VH}^*] & E[S_{HH}S_{VV}^*] \\ E[S_{VH}S_{HH}^*] & E[S_{VH}S_{VH}^*] & E[S_{VH}S_{VV}^*] \\ E[S_{VV}S_{HH}^*] & E[S_{VV}S_{VH}^*] & E[S_{VV}S_{VV}^*] \end{bmatrix} \tag{2.20}$$

é Hermitiana e definida positiva, em que  $E[\cdot]$  representa o operador de esperança estatística.

### 2.5.1 Estatísticas Gaussianas

Para Deng et al. (2017), sob a suposição de que o speckle está totalmente desenvolvido, verificou-se experimentalmente que as estatísticas Gaussianas geralmente fornecem um bom ajuste aos dados SAR, especialmente em áreas naturais homogêneas. Devido ao fato da distribuição Gaussiana multivariada ser matematicamente tratável, é adequada para modelar os vetores de espalhamento quando a área imageada é quase homogênea, ou seja, a rugosidade da superfície é relativamente baixa, a resolução espacial é moderada e um grande número

de espalhadores está presente. A suposição Gaussiana garante que as propriedades estatísticas dos dados são determinadas pela matriz de covariância  $\Sigma$ , onde se concentra toda informação. Então, a covariância amostral  $\mathbf{Z}$ , neste caso, segue uma distribuição Wishart complexa.

Quando o radar ilumina a área de uma superfície aleatória contendo muitos espalhadores elementares, como representado em (2.19),  $\mathbf{U}$ , pode ser modelado como tendo uma distribuição Gaussiana complexa multivariada, com média zero e covariância  $\Sigma$ , com Função Densidade de Probabilidade (PDF) (LEE et al., 1994):

$$f(\mathbf{U}; \Sigma) = \frac{1}{\pi^q |\Sigma|} \exp(-\mathbf{U}^* \Sigma^{-1} \mathbf{U}), \quad (2.21)$$

em que,  $|\cdot|$  é o determinante da matriz  $\Sigma$ .

Para facilitar a notação, Lee e Pottier (2009) trabalharam a partir da distribuição da matriz  $\mathbf{A} = L\mathbf{Z}$ , com  $\mathbf{Z}$  representada em (2.18). A matriz  $\mathbf{A}$  tem um Distribuição Wishart Complexa com função densidade de probabilidade (GOODMAN, 1963):

$$f_{\mathbf{A}}^{(L)}(\mathbf{A}; \Sigma, L) = \frac{|\mathbf{A}|^{L-q}}{\Gamma_q(L) |\Sigma|^L} \exp[-L \operatorname{tr}(\Sigma^{-1} \mathbf{A})], \quad (2.22)$$

em que,  $\operatorname{tr}(\cdot)$  é o operador traço,  $q$  é a dimensão do vetor  $\mathbf{U}$  e

$$\Gamma_q(L) = \pi^{q(q-1)/2} \prod_{i=1}^{q-1} \Gamma(L-i), \quad L \geq q \quad (2.23)$$

representa a função gama multivariada (constante de normalização).

Em conformidade com (2.22), o processo multilook definido em (2.18), segue uma Distribuição Wishart Complexa Escalonada (DWCE) (NASCIMENTO; FRERY; CINTRA, 2019), com função densidade de probabilidade, dada por:

$$f_{\mathbf{Z}}(\mathbf{Z}; \Sigma, L) = \frac{L^q |\mathbf{Z}|^{L-q}}{|\Sigma|^L \Gamma_q(L)} \exp[-L \operatorname{tr}(\Sigma^{-1} \mathbf{Z})]. \quad (2.24)$$

Portanto,  $\mathbf{Z}$  segue uma DWCE, em outras palavras,  $\mathbf{Z} \sim W_q(\Sigma, L)$ , satisfazendo (2.20), ou seja,  $E[\mathbf{Z}] = \Sigma$ , que é uma matriz auto-adjunta ou Hermitiana (uma matriz quadrada com entradas complexas cuja transposta de sua conjugada é igual à própria matriz) positiva definida (seus autovalores são todos positivos), então,  $\Sigma = \Sigma^*$ .

### 2.5.2 Coeficiente de correlação complexo

O coeficiente de correlação complexo é o parâmetro mais importante na descrição da função densidade de probabilidade, o qual, não é conhecido a priori e deve ser estimado utilizando um

número finito de amostras, e pode ser definido como:

$$\rho_c = \frac{E[S_i S_j^*]}{\sqrt{E[|S_i|^2]E[|S_j|^2]}} = |\rho_c| e^{i\theta}, \quad (2.25)$$

onde,  $0 < |\rho_c| < 1$  é a amplitude<sup>2</sup> do coeficiente de correlação,  $\theta$  é o ângulo de fase entre os canais. Além disso, o produto  $S_i S_j^*$  pode ser definido como,

$$S_i S_j^* = |S_i S_j^*| \exp(j(\theta_i - \theta_j)) = |S_i S_j^*| \exp(j\psi), \quad (2.26)$$

com,  $\{i, j\} \in \{HH, HV, VH, VV\}$ .

A coerência é uma fonte valiosa de informações nos dados PolSAR, pois, ao ser afetada por fatores como as estruturas dos alvos e a orientação da iluminação, ela revela características físicas dos elementos presentes na cena. Teoricamente, a coerência pode ser estimada usando duas intensidades multilook,  $Z_{ii}$  e  $Z_{jj}$  da matriz  $\mathbf{Z}$ , dada em (2.18). O coeficiente de correlação entre duas intensidades multilook é definido por:

$$\rho_I^{(L)} = \frac{E[(Z_{ii} - \bar{Z}_{ii})E(Z_{jj} - \bar{Z}_{jj})]}{\sqrt{E[(Z_{ii} - \bar{Z}_{ii})^2]E[(Z_{jj} - \bar{Z}_{jj})^2]}} \quad (2.27)$$

onde,  $\bar{Z}_{ii}$  e  $\bar{Z}_{jj}$  são obtidos das vizinhanças do pixel.

Para dados multilook, Lee et al. (1994) derivam a equivalência entre o coeficiente de correlação complexo e o correspondente às intensidades, tal relação é dada por:

$$\rho_I^{(L)} = |\rho_c|^2, \quad (2.28)$$

essa abordagem mostra que,  $\rho_I^{(L)}$  é independente do número de looks e, por outro lado, fornece uma maneira de calcular a coerência polarimétrica a partir do coeficiente de correlação entre intensidades. A estimativa do coeficiente de correlação a partir dos dados, não é uma simples média, e a estimativa da coerência usando (2.28) pode causar sérios problemas, já que os pixels vizinhos não são necessariamente homogêneos (LEE; POTTIER, 2009). Ou seja, o coeficiente de correlação varia com o tipo de cobertura da cena.

<sup>2</sup> A amplitude é a medida da força do sinal refletido pelo alvo e registrada pelo sistema. Além disso, ela está diretamente relacionada à distância do alvo, à rugosidade da superfície e à polarização do sinal.

### 3 METODOLOGIA

Neste capítulo, são apresentados os conceitos e definições de distribuições de probabilidade para dados circulares, abrangendo tanto modelos físicos quanto empíricos. Cada modelo é detalhado, destacando suas principais características, parâmetros relevantes e as inter-relações entre eles. Além disso, são discutidos os métodos de filtragem aplicados a dados interferométricos. Figuras foram incluídas ao longo do capítulo para ilustrar os conceitos abordados, contribuindo para uma melhor compreensão dos modelos e técnicas descritas.

#### 3.1 MODELOS PROBABILÍSTICOS PARA DADOS CIRCULARES

Os dados circulares surgem em uma grande variedade de campos de pesquisa (CREMERS; KLUGKIST, 2018), abrangendo desde a ecologia, onde se investigam informações sobre a direção do movimento dos animais, até as ciências médicas (MARDIA; TAYLOR; SUBRAMANIAM, 2006), onde a estrutura das proteínas ou a atividade neuronal é estudada por meio de medições periódicas, gerando dados circulares. Existem várias maneiras de representar uma observação circular, no entanto, é comum especificar uma distribuição circular como sendo a de um ângulo aleatório, o qual denotaremos por  $\Theta$ , correspondente a um ponto na circunferência do círculo unitário (ARTHUR; MARKUS; GRAEME, 2013).

Para lidar com esse tipo de dado, é necessário utilizar modelos probabilísticos que levem em conta essa estrutura cíclica e permitam à análise e interpretação dos resultados. Na modelagem estatística, os modelos podem ser categorizados como empíricos ou físicos, onde cada categoria pode ser classificada como univariada ou multivariada. A seguir apresentamos modelos estatísticos físicos e empíricos univariados, que tratam de dados circulares.

##### 3.1.1 Modelos circulares empíricos

Existem vários métodos gerais que podem ser usados para obter distribuições circulares. Segundo Arthur, Markus e Graeme (2013), talvez o mais simples seja a perturbação (perturbation). Nessa abordagem, a densidade de uma distribuição circular existente é multiplicada por alguma função escolhida de forma que o produto também seja uma densidade circular. Como exemplo desse tipo de construção, podemos destacar o modelo cardioide apresentado

na Subseção 3.1.1.2.

Outra abordagem clássica para obter distribuições circulares é o embrulhamento (wrapping). A definição por embrulhamento consiste em adaptar uma distribuição contínua com suporte definido na reta real para o intervalo  $(-\pi, \pi]$  ou  $[0, 2\pi)$ . Essa abordagem transforma a variável linear em circular por meio da operação módulo  $2\pi$ , assegurando a periodicidade da densidade resultante (ver Apêndice B.1). Os modelos Normal wrapped e Cauchy wrapped, das Subseções 3.1.1.3 e 3.1.1.4, são exemplos desse tipo de construção. Para mais informações sobre essa técnica, consultar (RAO; SENGUPTA, 2001) e (ARTHUR; MARKUS; GRAEME, 2013).

Em contraste com abordagens baseadas na adaptação de distribuições definidas na reta real, a construção direta de distribuições circulares busca formular modelos já definidos sobre o círculo, sem a necessidade de alterar o suporte original (ARTHUR; MARKUS; GRAEME, 2013). Como exemplos clássicos dessa abordagem, podemos citar a distribuição circular uniforme (Subseção 3.1.1.1) e a distribuição de von Mises (Subseção 3.1.1.5). Este último modelo foi introduzido por von Mises, em 1918, para estudar os desvios dos pesos atômicos medidos em relação a valores inteiros. De acordo com Mardia (1975), o modelo von Mises desempenha um papel fundamental na inferência estatística no círculo, tendo uma importância análoga à da distribuição normal.

### 3.1.1.1 Distribuição Uniforme Circular

Seja  $\Theta$  uma Variável Aleatória Circular ( $\mathcal{VAC}$ ) tal que sua função de densidade (RAO; SENGUPTA, 2001) é dada por:

$$f_{\Theta}(\theta) = \frac{1}{2\pi} \text{ para } -\pi < \theta \leq \pi. \quad (3.1)$$

Neste caso,  $\Theta$  tem Distribuição Uniforme Circular ( $\mathcal{DUC}$ ) no intervalo  $(-\pi, \pi]$ . Observamos que a direção média (ver Apêndice B.1) de uma variável aleatória com  $\mathcal{DUC}$  não está bem definida, no entanto, não deixa de apresentar relevância quando se analisa os dados circulares, pois ela retrata a ausência da direção preferencial e o comprimento do vetor resultante (ver Apêndice B.1) é sempre igual a zero. Integrando (3.1), a Função de Distribuição Acumulada (CDF) de  $\Theta$  é simplesmente,

$$F_{\Theta}(\theta) = \frac{\theta + \pi}{2\pi} \text{ para } -\pi < \theta \leq \pi. \quad (3.2)$$

### 3.1.1.2 Distribuição Cardioide

Seja  $\Theta$  uma  $\mathcal{VAC}$ . A Distribuição Cardioide ( $\mathcal{DC}$ ) é derivada da curva cardioide e tem função densidade de probabilidade (RAO; SENGUPTA, 2001), (ARTHUR; MARKUS; GRAEME, 2013), (PAULA et al., 2021), definida por:

$$f_{\Theta}(\theta; \mu, \rho) = \frac{1}{2\pi} \{1 + 2\rho \cos(\theta - \mu)\} \text{ para } -\pi < \theta \leq \pi. \quad (3.3)$$

em que,  $-\pi < \mu \leq \pi$  é a direção média e  $|\rho| \leq 0,5$  o parâmetro de concentração, ou seja, indica a dispersão dos dados circulares em torno da direção média. A  $\mathcal{DC}$  é unimodal e simétrica em torno da média  $\mu$ . Integrando o modelo em (3.3), obtemos a CDF, definida por:

$$F_{\Theta}(\theta; \mu, \rho) = \frac{\theta + \pi}{2\pi} + \frac{\rho}{\pi} [(\sin(\theta - \mu) - \sin \mu)] \text{ para } -\pi < \theta \leq \pi. \quad (3.4)$$

### 3.1.1.3 Distribuição normal wrapped

A distribuição Normal Wrapped ( $\mathcal{WN}$ ) é simétrica com dois parâmetros, podendo ser obtida envolvendo a distribuição Normal Padrão ou Gaussiana com média  $\mu$  e variância  $\sigma^2$ , em torno do círculo unitário (ver apêndice B.3). Ou seja, se uma variável aleatória  $X$  tem distribuição Normal, então a variável aleatória circular  $\Theta$  tem distribuição  $\mathcal{WN}$  com parâmetros  $\mu$  e  $\sigma^2$ . Assim, para a variável  $\Theta$ , a função densidade de probabilidade da distribuição  $\mathcal{WN}$  (RAO; SENGUPTA, 2001; ARTHUR; MARKUS; GRAEME, 2013; KURZ; GILITSCHENSKI; HANEBECK, 2014) é expressa por:

$$\begin{aligned} f_{\Theta}(\theta; \mu, \sigma^2) &= \sum_{k=-\infty}^{\infty} g(\theta + 2k\pi) \\ &= \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \exp\left[-\frac{(\theta - \mu + 2k\pi)^2}{2\sigma^2}\right] \text{ para } -\pi < \theta \leq \pi, \end{aligned} \quad (3.5)$$

com parâmetro de localização  $\mu \in (-\pi, \pi]$  e parâmetro de escala  $\sigma > 0$ .

Os termos da série em (3.5) decaem exponencialmente e tendem a zero. Isso ocorre porque, devido à natureza da distribuição Gaussiana, a contribuição de cada termo diminui rapidamente à medida que  $k$  aumenta. Dessa forma, para valores elevados de  $k$ , os termos tornam-se praticamente nulos, o que permite aproximar a função densidade de probabilidade utilizando uma série truncada (KURZ; GILITSCHENSKI; HANEBECK, 2014),

$$\begin{aligned} f_{\Theta}(\theta; \mu, \sigma^2) &\approx f_n(\theta; \mu, \sigma^2) \\ &= \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-n}^n \exp\left[-\frac{(\theta - \mu + 2k\pi)^2}{2\sigma^2}\right] \text{ para } -\pi < \theta \leq \pi, \end{aligned} \quad (3.6)$$

em que, apenas  $2n + 1$  termos da soma são considerados.

Uma forma alternativa e mais útil de representar a densidade em (3.6) é apresentada por Arthur, Markus e Graeme (2013), Kurz, Gilitschenski e Hanebeck (2014) e expressa por:

$$g_{\Theta}(\theta; \mu, \rho) = \frac{1}{2\pi} \left( 1 + 2 \sum_{k=1}^{\infty} \rho^{k^2} \cos k(\theta - \mu) \right) \text{ para } -\pi < \theta \leq \pi. \quad (3.7)$$

onde  $\rho = \exp(-\sigma^2/2)$ . Análogo a (3.6), definimos uma versão truncada da densidade em (3.7), definida por:

$$\begin{aligned} g_{\Theta}(\theta; \mu, \sigma^2) &\approx g_n(\theta; \mu, \sigma^2) \\ &= \frac{1}{2\pi} \left( 1 + 2 \sum_{k=1}^n \rho^{k^2} \cos k(\theta - \mu) \right) \text{ para } -\pi < \theta \leq \pi. \end{aligned} \quad (3.8)$$

O parâmetro  $n$ , tanto em  $f_n$  quanto em  $g_n$ , é tratado da mesma maneira. No entanto, a avaliação de  $f_n$  envolve  $2n + 1$  termos da soma, enquanto a avaliação de  $g_n$  envolve apenas  $n$  termos da soma (KURZ; GILITSCHENSKI; HANEBECK, 2014). Observamos que, quando  $\rho \rightarrow 0$ , a  $\mathcal{WN}$  converge para a distribuição uniforme circular.

#### 3.1.1.4 Distribuição de Cauchy wrapped

A distribuição Cauchy Wrapped ( $\mathcal{WN}$ ) é obtida embrulhando a distribuição Cauchy Padrão em torno do círculo unitário, transformando os valores da distribuição original, que podem ser definidos ao longo da reta real, para o intervalo  $(-\pi, \pi]$ . Na reta, a distribuição Cauchy Padrão tem função densidade de probabilidade dada por:

$$g(x) = \left( \frac{1}{\pi} \right) \frac{\sigma}{\sigma^2 + (x - \mu)^2} \text{ para } \sigma > 0 \text{ e } -\infty < x < \infty,$$

onde  $\mu$  é a mediana da distribuição e  $\sigma$  é o parâmetro de escala.

Seja  $\Theta$  uma  $\mathcal{VAC}$ . De acordo com Rao e SenGupta (2001), Arthur, Markus e Graeme (2013), a função densidade de probabilidade da distribuição  $\mathcal{WC}$  é dada da seguinte forma,

$$\begin{aligned} f_{\Theta}(\theta; \mu, \rho) &= \frac{1}{2\pi} \left( 1 + 2 \sum_{k=1}^{\infty} \rho^k \cos k(\theta - \mu) \right) \\ &= \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)} \text{ para } -\pi < \theta \leq \pi. \end{aligned} \quad (3.9)$$

em que,  $\rho = e^{-\sigma} \in [0, 1]$ . A igualdade das duas expressões acima é verificada pela equação das partes reais da identidade da série geométrica (RAO; SENGUPTA, 2001) (ver Apêndice B.6),

$$\sum_{k=1}^{\infty} a^k = \frac{a}{1 - a},$$

com  $a = \rho e^{-i(\theta-\mu)}$ .

Pela propriedade (3) das distribuição wrapped no Apêndice B, temos:

$$f_{\Theta}(\theta; \mu, \rho) = \sum_{k=-\infty}^{\infty} g(\theta + 2k\pi) = \frac{1}{2\pi} \left( 1 + 2 \sum_{k=1}^{\infty} (\alpha_k \cos(k\theta) + \beta_k \sin(k\theta)) \right). \quad (3.10)$$

Substituindo  $\alpha_k = e^{-\sigma|k|} \cos(k\mu)$  e  $\beta_k = e^{-\sigma|k|} \sin(k\mu)$  em (3.10) e considerando  $\rho = e^{-\sigma}$ , obtemos:

$$\begin{aligned} f_{\Theta}(\theta; \mu, \rho) &= \frac{1}{2\pi} \left( 1 + 2 \sum_{k=-\infty}^{\infty} e^{-\sigma|k|} \cos k(\theta - \mu) \right) \\ &= \frac{1}{2\pi} \left( 1 + 2 \sum_{k=-\infty}^{\infty} \rho^{|k|} \cos k(\theta - \mu) \right) \text{ para } -\pi < \theta \leq \pi. \end{aligned} \quad (3.11)$$

Como observado anteriormente para a distribuição  $\mathcal{WN}$ , quando  $\rho \rightarrow 0$ , a  $\mathcal{WC}$  converge para a uniforme circular (ARTHUR; MARKUS; GRAEME, 2013).

### 3.1.1.5 Distribuição de von Mises

Seja  $\Theta$  uma  $\mathcal{VAC}$ . A distribuição von Mises ( $\mathcal{VM}$ ) ou Normal Circular tem função densidade (RAO; SENGUPTA, 2001), (ARTHUR; MARKUS; GRAEME, 2013), dada por:

$$f_{\Theta}(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta-\mu)}, \text{ para } -\pi < \theta \leq \pi, \quad (3.12)$$

onde,  $\kappa \geq 0$  é o parâmetro de concentração,  $-\pi < \mu \leq \pi$  é o parâmetro de locação ou ângulo médio, e  $I_0(\kappa)$  a função de Bessel modificada de primeiro tipo de ordem zero (KATO; JONES, 2010) (ver Apêndice A), a qual é definida por:

$$I_0(\kappa) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\kappa \cos \theta} d\theta = \sum_{r=0}^{\infty} \left( \frac{\kappa}{2} \right)^{2r} \left( \frac{1}{r!} \right)^2. \quad (3.13)$$

Neste modelo, podemos observar que, quando  $\kappa = 0$ , a  $\mathcal{VM}$  se reduz a distribuição uniforme circular no intervalo  $(-\pi, \pi]$ , qualquer que seja o valor de  $\mu$ . A distribuição  $\mathcal{VM}(\mu, \kappa)$  satisfaz outras propriedades bastante interessantes, dentre elas, evidenciamos:

1. Quando  $\kappa \rightarrow \infty$ , a  $\mathcal{VM}$  converge para a distribuição degenerada no ponto  $\theta = \mu$ ;
2. Para valores não muito pequenos de  $\kappa$ , a  $\mathcal{VM}$  se aproxima de uma distribuição normal com parâmetros  $\mu$  e  $1/\kappa$ .

Para mais informações sobre as propriedades da distribuição von Mises, consultar (MARDIA K. V. AND JUPP, 2000; RAO; SENGUPTA, 2001; KATO; JONES, 2010) e (ARTHUR; MARKUS; GRAEME, 2013).

### 3.1.2 Modelos físicos para diferença de fases

Para dados PolSAR multilook, as informações polarimétricas, como amplitude e fase, são geradas e obtidas pela emissão de sinais complexos multivariados. Em sensoriamento remoto por radar, o conceito de fase é aplicado às oscilações de ondas eletromagnéticas. Além disso, a fase se caracteriza por seu aspecto periódico e pode ser definida como o tempo de viagem bidirecional do sinal SAR entre os sensores e o solo, incluindo toda a propagação e atrasos inerentes às dispersões. Cada ângulo de fase é distribuído uniformemente no intervalo  $(-\pi, \pi]$  e não contém informações sobre as propriedades da cena. No entanto, a diferença de fase ou fase interferométrica pode conter informações detalhadas sobre as propriedades da cena observada, como relevância topográfica ou estrutural.

O objetivo em derivar expressões de forma fechada para a diferença de fase vem do fato de que, juntamente com o produto de duas amplitudes, reflete a correlação entre diferentes canais polarimétricos (DENG et al., 2017). Neste sentido, Lee et al. (1994) derivam uma expressão para a diferença de fase entre qualquer par de componentes polarizados. Segundo Lee et al. (1994), a diferença de fase single look, pode ser definida como:

$$\psi_1 = \angle S_i S_j^*. \quad (3.14)$$

A fase multilook é obtida pela expressão apresentada em 3.15. É importante destacar que o multilook não equivale à média das diferenças de fase single look.

$$\psi_L = \angle \left[ \frac{1}{L} \sum_{\ell=1}^L S_i(\ell) S_j^*(\ell) \right], \quad (3.15)$$

em que,  $\psi_L$  é o argumento de um termo fora da diagonal principal da matriz de covariância amostral  $\mathbf{Z}$ ,  $L$  é o número de looks e o símbolo  $\angle$  representa o operador angular<sup>1</sup>. Por conveniência na notação, o subscrito  $L$  será omitido.

Visto que as densidades envolvem apenas duas polarizações, (LEE et al., 1994) utilizam a distribuição em (2.22), e para  $q = 2$ , podemos escrever:

$$\mathbf{A} = \begin{bmatrix} A_{11} & \alpha \exp(j\psi) \\ \alpha \exp(-j\psi) & A_{22} \end{bmatrix} \quad (3.16)$$

<sup>1</sup> Na expressão em (3.14) o operador  $\angle$  denota o ângulo (ou fase) de um número complexo formado pelos dois números complexos conjugados  $S_i$  e  $S_j^*$  no plano complexo. Enquanto, na expressão em (3.15) retorna o ângulo do valor complexo avaliado na matriz de covariância complexa hermitiana  $\mathbf{Z}$ , ou seja, calcula o ângulo formado pelo valor complexo de um elemento fora da diagonal principal da matriz hermitiana no plano complexo.

com sua matriz de covariância, dada por:

$$\mathbf{\Sigma} = \begin{bmatrix} \Sigma_{11} & \sqrt{\Sigma_{11}\Sigma_{22}}|\rho_c| \exp(j\theta) \\ \sqrt{\Sigma_{11}\Sigma_{22}}|\rho_c| \exp(-j\theta) & \Sigma_{22} \end{bmatrix} \quad (3.17)$$

em que,  $\alpha \exp(j\psi) = A_{12R} + jA_{12I}$  e  $\Sigma_{ii} = E(|S_i|^2)$ .

Por conveniência, normalizamos as intensidades de  $A_{11}$  e  $A_{22}$ , e  $\alpha$  em termos dos elementos da matriz  $\mathbf{\Sigma}$ , por:

$$B_1 = \frac{A_{11}}{\Sigma_{11}}, \quad B_2 = \frac{A_{22}}{\Sigma_{22}}, \quad \eta = \frac{\alpha}{\sqrt{\Sigma_{11}\Sigma_{22}}}. \quad (3.18)$$

Agora, substituindo o vetor  $[A_{11}, A_{22}, A_{12R}, A_{12I}]$  por  $[B_1, B_2, \eta, \psi]$ , respectivamente em (2.22), obtemos:

$$f_{\mathbf{A}}(B_1, B_2, \eta, \psi; \theta, \rho_c, L) = \frac{(B_1 B_2 - \eta^2)^{L-2} \eta}{\pi (1 - |\rho_c|^2)^L \Gamma(L) \Gamma(L-1)} \times \exp\left(-\frac{B_1 + B_2 - 2\eta|\rho_c| \cos(\psi - \theta)}{(1 - |\rho_c|^2)}\right). \quad (3.19)$$

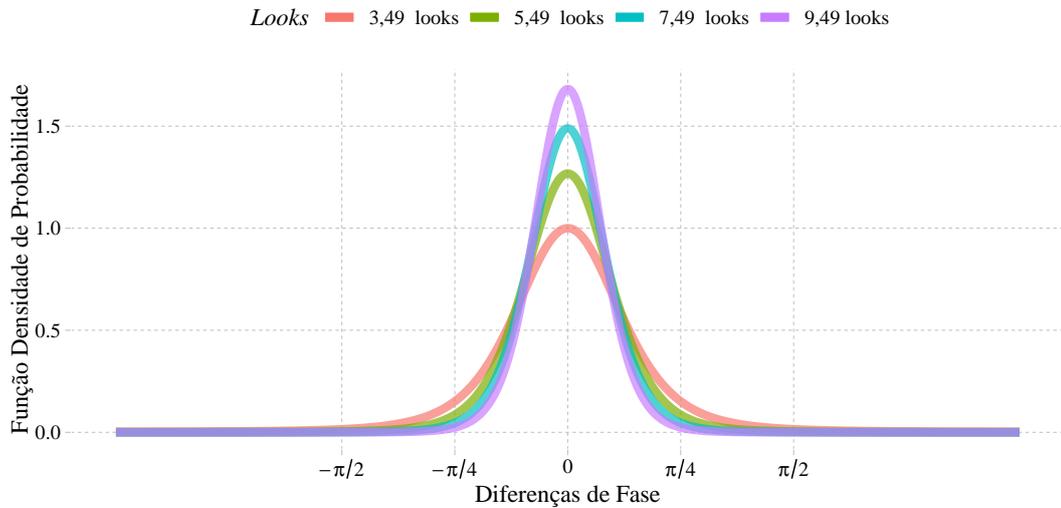
Integrando a função em (3.19) com relação as variáveis  $B_1, B_2$  e  $\eta$  com domínio de integração restrito por  $B_1 B_2 - \eta^2 > 0$ , a função densidade para diferença de fase multilook é expressa da seguinte forma:

$$f(\psi; \theta, \rho_c, L) = \frac{\Gamma(L + \frac{1}{2})(1 - |\rho_c|^2)^L \beta}{2\sqrt{\pi}\Gamma(L)(1 - \beta^2)^{L+\frac{1}{2}}} + \frac{(1 - |\rho_c|^2)^L}{2\pi} {}_2F_1\left(L, 1; \frac{1}{2}; \beta^2\right), \quad (3.20)$$

em que,  $-\pi < \psi \leq \pi$  é a fase,  $\rho_c$  é o coeficiente de correlação complexa,  $\beta = |\rho_c| \cos(\psi - \theta)$ ,  $\theta$  é a fase entre os canais, e  ${}_2F_1$  é a Função Hipergeométrica de Gauss (GHF). A Figura 8, ilustra a expressão em (3.20) para  $L = 3, 49; 5, 49; 7, 49$  e  $9, 49$ , com  $|\rho_c| = 0,7$  e  $\theta = 0$ .

Devido à dependência da GHF, a computação e a análise numérica das propriedades estatísticas da expressão em (3.20), pode ser extremamente difícil ou intratável para um  $L$  grande e/ou valores de correlação elevados (HAYNES, 2017). Para resolver este problema, Gierull (2021) propôs uma nova forma para o modelo em (3.20). Ao restringir o número de looks a inteiros positivos maiores que 1, essa restrição torna o modelo numericamente tratável. A

Figura 8 – Múltiplas visadas para distribuição de diferença de fase: Jong-Sen Lee



Fonte: Lee et al. (1994)

PDF torna-se:

$$f(\psi; \theta, \rho_c, L) = \frac{(1 - |\rho_c|^2)^L}{2\pi\Gamma(L)(1 - \beta^2)^{L+\frac{1}{2}}} \left\{ \sqrt{\pi}\Gamma\left(L + \frac{1}{2}\right)\beta \right. \\ \left. + \frac{\Gamma\left(L - \frac{1}{2}\right)}{\sqrt{\pi}} \left[ \sqrt{1 - \beta^2} + 2\left(L - \frac{1}{2}\right)\beta \arcsen \beta \right] \right. \\ \left. + \sum_{k=1}^{L-1} \frac{(-1)^k}{2} \frac{\Gamma\left(k - L + \frac{1}{2}\right)\Gamma(L - k)}{\Gamma\left(\frac{3}{2} - L\right)} \frac{1 + (2k - 1)\beta^2}{(1 - \beta^2)^{k-L+\frac{1}{2}}} \right\}. \quad (3.21)$$

em que,  $-\pi < \psi \leq \pi$ .

O modelo em (3.21) omite a GHF e seus problemas numéricos, no entanto, envolve a função gama, que pode apresentar desafios numéricos significativos. Gierull (2021) discute estratégias para contornar as dificuldades relacionadas às razões de grandes quantidades derivadas das funções gama. Para resolver esses problemas, (GIERULL, 2021) utiliza a propriedade da função log-gama, definida em (3.22):

$$\log \left[ \frac{\Gamma(x + k)}{\Gamma(x + 1)} \right] = \sum_{i=1}^{k-1} (x + k - i) \text{ para } x = \frac{1}{2} - L. \quad (3.22)$$

De acordo com Gierull (2021), os problemas de instabilidade numérica atribuídos a (3.21) estão relacionados à dificuldade de calcular diretamente a soma para valores grandes de  $L$  e/ou coerência próxima de 1. Esses problemas são superados utilizando a função log-gama, que permite somar ou subtrair grandes exponenciais, em vez de multiplicá-las ou dividi-las.

Essa abordagem oferece uma solução elegante para evitar problemas numéricos ao calcular produtos ou divisões de grandes quantidades. Por exemplo,  $\text{gamma}(1000)/\text{gamma}(999)$

no R resulta em NaN. Por outro lado, ao usar  $\exp(\lgamma(1000)-\lgamma(999))$  produz 999. Então matematicamente a razão  $A/B$  é simplesmente expressa como  $\exp(\ln(A/B)) = \exp(\ln(A)-\ln(B))$ , que é computável quando  $\ln(A)-\ln(B)$  se torna um valor razoável. Além disso,  $\lgamma(N)$  e  $\gamma(\ln(N))$  podem ser calculados numericamente de forma eficiente sem avaliar a função gama em si. Esses foram os motivos para usar as funções  $\ln$  e  $\exp$  na implementação do modelo em (3.21) no R.

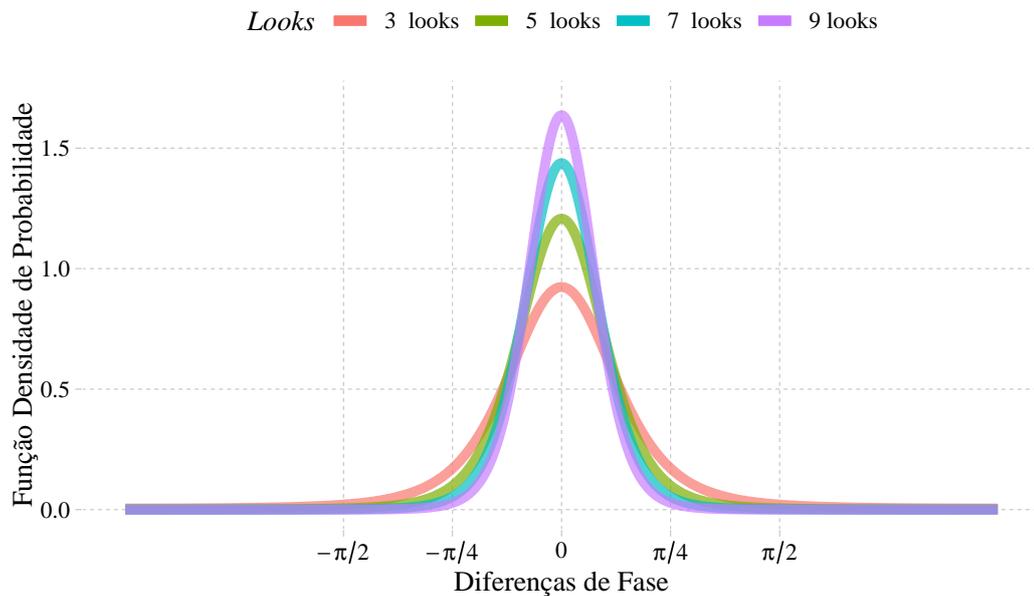
Soluções de forma fechada para a expressão em (3.20), quando  $L$  é um número inteiro positivo, foram fornecidas até a ordem 4 ( $L = 1; 2; 3; 4$ ) por Lee et al. (1994), usando as fórmulas de recorrências de Gauss. Para um número maior de looks, essas soluções estão se tornando longas e complicadas, na melhor das hipóteses. Esses desafios tem motivado vários autores a investigar soluções aproximadas (SINTES et al., 2012) ou mais recentemente (HAYNES, 2017) e (GIERULL, 2021).

Para  $L$  inteiro positivo e maior do que 1, as expressões em (3.20) e (3.21) são idênticas. Além disso, para  $L = 2$  esses modelos têm a mesma expressão, definida por:

$$f(\psi; \theta, \rho_c, 2) = \frac{3(1 - |\rho_c|^2)^2 \beta}{8(1 - \beta^2)^{\frac{5}{2}}} + \frac{(1 - |\rho_c|^2)^2}{4\pi(1 - \beta^2)^2} \cdot \left( 1 + (1 + \beta^2) + \frac{3\beta \arcsen \beta}{\sqrt{1 - \beta^2}} \right). \quad (3.23)$$

A Figura 9, apresenta a expressão em (3.21), adotando  $L = 3; 5; 7$  e  $9$ ,  $|\rho_c| = 0,7$  e  $\theta = 0$ . Visualmente, observamos uma pequena diferença ao comparar as Figuras 8 e 9 para valores de  $L$  aproximados, indicando que não há perda significativa de informações ao utilizar o modelo (3.21) em comparação com o modelo (3.20).

Figura 9 – Múltiplas visadas para distribuição de diferença de fase: Christoph H. Gierull



Fonte: O autor (2025).

### 3.2 MODELOS EMPÍRICOS TRUNCADOS

Nesta seção, apresentamos as distribuições truncadas: Normal e Cauchy. A escolha desses modelos é justificada por sua ampla aplicação em problemas práticos, especialmente o modelo Normal, que é frequentemente adotado por pesquisadores em estudos estatísticos devido à sua versatilidade e relevância em diversas áreas. Os modelos truncados discutidos nesta seção surgem ao limitar o domínio de uma distribuição existente, resultando em modelos truncados a partir dessas limitações.

Restringir o suporte dos modelos originais consiste em limitar a densidade a um intervalo específico  $(a, b]$ , utilizando a função indicadora  $\mathbb{1}_{(a,b]}(\cdot)$ . Essa função garante que a densidade seja nula fora do intervalo truncado, permanecendo diferente de zero apenas dentro dos limites definidos. Essa abordagem contrasta com os modelos wrapped, nos quais todos os valores da reta real são mapeados para a circunferência de raio 1 por meio da operação módulo  $2\pi$ .

Após o truncamento, a função de densidade precisa ser ajustada para garantir que continue válida, ou seja, que sua integral sobre o novo domínio seja igual a 1. Para isso, a densidade original é normalizada dividindo-se pelo valor da função de distribuição acumulada no intervalo  $(a, b]$ . Dessa forma, asseguramos que a nova função de densidade mantenha a propriedade fundamental de que a área sob a curva em seu domínio de definição corresponda à totalidade das probabilidades.

**Definição 1.** Seja  $Z$  uma variável aleatória com função densidade de probabilidade  $\phi$ , pertencente a uma família de distribuições padronizadas, e seja  $G$  sua Função de Distribuição Acumulada, definida por:

$$G(z) = \int_{-\infty}^z \phi(t) dt. \quad (3.24)$$

Se  $Z$  está restrito ao intervalo fixo e conhecido  $(a, b)$ , com  $-\infty < a < b < \infty$ , então a variável aleatória truncada  $Z_t = Z \mid (a < Z < b)$  possui densidade de probabilidade dada por (NADARAJAH; KOTZ, 2006a):

$$f_{Z_t}(z; a, b) = \begin{cases} \frac{\phi(z)}{G(b) - G(a)}, & \text{se } a < z \leq b, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.25)$$

Ao aplicarmos uma transformação de escala e localização sobre a variável truncada  $Z_t$ , obtemos uma nova variável aleatória  $Y$  definida por:

$$Y = \mu + \sigma Z_t, \quad (3.26)$$

com  $\mu \in \mathbb{R}$  e  $\sigma > 0$ . De acordo com Johnson, Kotz e Balakrishnan (1994) e Robert (1995), a função densidade de probabilidade de  $Y$ , truncada bilateralmente no intervalo  $(a, b]$ , é dada por:

$$f_Y(y; \mu, \sigma, a, b) = \frac{\phi\left(\frac{y - \mu}{\sigma}\right)}{\sigma \left[ G\left(\frac{b - \mu}{\sigma}\right) - G\left(\frac{a - \mu}{\sigma}\right) \right]} \mathbb{1}_{(a, b]}(y), \quad (3.27)$$

em que a função indicadora,  $\mathbb{1}_{(a, b]}(y)$ , assegura que o suporte da densidade esteja contido no intervalo truncado  $(a, b]$ .

A partir do modelo (3.27), ao considerarmos os limites  $a = -\pi$  e  $b = \pi$ , podemos construir um modelo truncado com base em uma função de densidade de probabilidade com suporte não limitado,  $\phi: \mathbb{R} \rightarrow \mathbb{R}_+$ . Esse modelo consiste em restringir a função ao intervalo  $(-\pi, \pi]$  e aplicar uma normalização apropriada, de forma que a integral da densidade no intervalo truncado seja igual a 1. Usando as seguintes densidades para  $\phi$ ,

$$\phi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \quad \text{ou} \quad \phi(y) = \frac{1}{\pi(1 + y^2)},$$

na expressão (3.27), obtemos a Normal Truncada  $\mathcal{NT}(\mu, \sigma)$  e Cauchy Truncada  $\mathcal{CT}(\mu, \sigma)$ , respectivamente. Nadarajah e Kotz (2006a) apresenta funções em R para o cálculo de PDF, CDF, função de quantis, geração de amostras aleatórias, além do cálculo da esperança e variância, para qualquer distribuição nativamente implementada no R.

### 3.3 FUNÇÃO DE LOG-VEROSSIMILHANÇA

Nesta seção, exibimos a função de log-verossimilhança para os modelos físicos em (3.20) e (3.21), conforme propostos por Lee et al. (1994) e Gierull (2021), respectivamente.

#### 3.3.1 Função de log-verossimilhança do modelo definido em (3.20)

A expressão de forma fechada em (3.20), apresentada por Lee et al. (1994) tem a seguinte função de densidade para diferença de fase:

$$f(\psi; \theta, \rho_c, L) = \frac{\Gamma(L + \frac{1}{2})(1 - |\rho_c|^2)^L \beta}{2\sqrt{\pi}\Gamma(L)(1 - \beta^2)^{L+\frac{1}{2}}} + \frac{(1 - |\rho_c|^2)^L}{2\pi} {}_2F_1\left(L, 1; \frac{1}{2}; \beta^2\right).$$

Aplicando o logaritmo natural em ambos os lados de (3.20), obtemos:

$$\begin{aligned} \ln f(\psi; \theta, \rho_c, L) &= \ln \left[ \frac{\Gamma(L + \frac{1}{2})(1 - |\rho_c|^2)^L \beta}{2\sqrt{\pi}\Gamma(L)(1 - \beta^2)^{L+\frac{1}{2}}} + \frac{(1 - |\rho_c|^2)^L}{2\pi} {}_2F_1\left(L, 1; \frac{1}{2}; \beta^2\right) \right] \\ &= \ln \left\{ \frac{(1 - |\rho_c|^2)^L}{2} \left[ \frac{\Gamma(L + \frac{1}{2})\beta}{\sqrt{\pi}\Gamma(L)(1 - \beta^2)^{L+\frac{1}{2}}} + \pi^{-1} {}_2F_1\left(L, 1; \frac{1}{2}; \beta^2\right) \right] \right\} \\ &= \ln \left[ \frac{(1 - |\rho_c|^2)^L}{2} \right] + \ln \left[ \frac{\Gamma(L + \frac{1}{2})\beta}{\sqrt{\pi}\Gamma(L)(1 - \beta^2)^{L+\frac{1}{2}}} + \pi^{-1} {}_2F_1\left(L, 1; \frac{1}{2}; \beta^2\right) \right]. \end{aligned}$$

Agora, reescrevendo a equação anterior, obtemos:

$$\begin{aligned} \ln f(\psi; \theta, \rho_c, L) &= \\ &L \ln(1 - |\rho_c|^2) - \ln 2 + \ln \left[ \frac{\Gamma(L + \frac{1}{2})\beta}{\sqrt{\pi}\Gamma(L)(1 - \beta^2)^{L+\frac{1}{2}}} + \pi^{-1} {}_2F_1\left(L, 1; \frac{1}{2}; \beta^2\right) \right]. \end{aligned}$$

Considerando  $\boldsymbol{\psi} = (\psi_1, \psi_2, \dots, \psi_n)$  os valores observados de uma determinada amostra, temos:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\psi}; \theta, \rho_c, L) &= \ln \prod_{i=1}^n f(\psi_i; \theta, \rho_c, L) \\ &= \sum_{i=1}^n \ln f(\psi_i; \theta, \rho_c, L). \end{aligned}$$

Portanto, a função de log-verossimilhança do modelo dada por Lee et al. (1994), tem a seguinte forma:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\psi}; \theta, \rho_c, L) &= nL \ln(1 - |\rho_c|^2) - n \ln 2 \\ &+ \sum_{i=1}^n \ln \left[ \frac{\Gamma(L + \frac{1}{2})|\rho_c| \cos(\psi_i - \theta)}{\sqrt{\pi}\Gamma(L)(1 - |\rho_c|^2 \cos^2(\psi_i - \theta))^{L+\frac{1}{2}}} + \pi^{-1} {}_2F_1\left(L, 1; \frac{1}{2}; |\rho_c|^2 \cos^2(\psi_i - \theta)\right) \right]. \end{aligned}$$

### 3.3.2 Função de log-verossimilhança do modelo definido em (3.21)

A função de densidade para diferença de fase em (3.21), apresentada por Gierull (2021) é definida por:

$$f(\psi; \theta, \rho_c, L) = \frac{(1 - |\rho_c|^2)^L}{2\pi\Gamma(L)(1 - \beta^2)^{L+\frac{1}{2}}} \left\{ \sqrt{\pi}\Gamma\left(L + \frac{1}{2}\right)\beta \right. \\ \left. + \frac{\Gamma\left(L - \frac{1}{2}\right)}{\sqrt{\pi}} \left[ \sqrt{1 - \beta^2} + 2\left(L - \frac{1}{2}\right)\beta \arcsen \beta \right] \right. \\ \left. + \sum_{k=1}^{L-1} \frac{(-1)^k}{2} \frac{\Gamma\left(k - L + \frac{1}{2}\right)\Gamma(L - k)}{\Gamma\left(\frac{3}{2} - L\right)} \frac{1 + (2k - 1)\beta^2}{(1 - \beta^2)^{k-L+\frac{1}{2}}} \right\}.$$

Fazendo,

$$A = \frac{(1 - |\rho_c|^2)^L}{2\pi\Gamma(L)(1 - \beta^2)^{L+\frac{1}{2}}}, \quad B = \sqrt{\pi}\Gamma\left(L + \frac{1}{2}\right)\beta, \quad C = \frac{\Gamma\left(L - \frac{1}{2}\right)}{\sqrt{\pi}}, \\ D = \sqrt{1 - \beta^2}, \quad E = 2\left(L - \frac{1}{2}\right)\beta \arcsen \beta, \quad F = \frac{(-1)^k}{2}, \\ G = \frac{\Gamma\left(k - L + \frac{1}{2}\right)\Gamma(L - k)}{\Gamma\left(\frac{3}{2} - L\right)}, \quad H = \frac{1 + (2k - 1)\beta^2}{(1 - \beta^2)^{k-L+\frac{1}{2}}},$$

podemos reescrever a expressão em (3.21) do seguinte modo:

$$f(\psi; \theta, \rho_c, L) = A[B + C(D + E) + FGH].$$

Aplicando o logaritmo natural, obtemos:

$$\begin{aligned} \ln f(\psi; \theta, \rho_c, L) &= \ln\{A[B + C(D + E) + FGH]\} \\ &= \ln A + \ln[B + C(D + E) + FGH] \\ &= \ln A + \ln(B + DC + CE + FGH) \end{aligned}$$

Considerando  $\psi = (\psi_1, \psi_2, \dots, \psi_n)$  os valores observados de uma amostra, a função de

log-verossimilhança da densidade dada por Gierull (2021) é definida por:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\psi}; \theta, \rho_c, L) &= \ln \prod_{i=1}^n f(\psi_i; \theta, \rho_c, L) = \sum_{i=1}^n \ln f(\psi_i; \theta, \rho_c, L) = \\ &= nL \ln(1 - |\rho_c|^2) - n \ln 2 - n \ln \pi - \ln \Gamma(L)n - \left(L + \frac{1}{2}\right) \sum_{i=1}^n \ln [1 - |\rho_c|^2 \cos^2(\psi_i - \theta)] \\ &+ \ln \left\{ \sqrt{\pi} \Gamma\left(L - \frac{1}{2}\right) |\rho_c| \sum_{i=1}^n \cos(\psi_i - \theta) + \frac{\Gamma\left(L - \frac{1}{2}\right)}{\sqrt{\pi}} \sum_{i=1}^n [1 - |\rho_c|^2 \cos^2(\psi_i - \theta)]^{\frac{1}{2}} \right. \\ &+ 2|\rho_c| \frac{\Gamma\left(L - \frac{1}{2}\right)}{\sqrt{\pi}} \left(L - \frac{1}{2}\right) \sum_{i=1}^n \cos(\psi_i - \theta) \sum_{i=1}^n \arcsen[|\rho_c|(\psi_i - \theta)] \\ &\left. + \sum_{i=1}^n \sum_{k=1}^{L-1} \frac{(-1)^k \Gamma\left(k - L + \frac{1}{2}\right) \Gamma(L - k)}{2 \Gamma\left(\frac{3}{2} - L\right)} \frac{1 + (2k - 1)[|\rho_c|^2 \cos(\psi_i - \theta)]}{[1 - |\rho_c|^2 \cos^2(\psi_i - \theta)]^{k-L+\frac{1}{2}}} \right\}. \end{aligned}$$

### 3.4 RELAÇÕES ENTRE MODELOS ESTATÍSTICOS

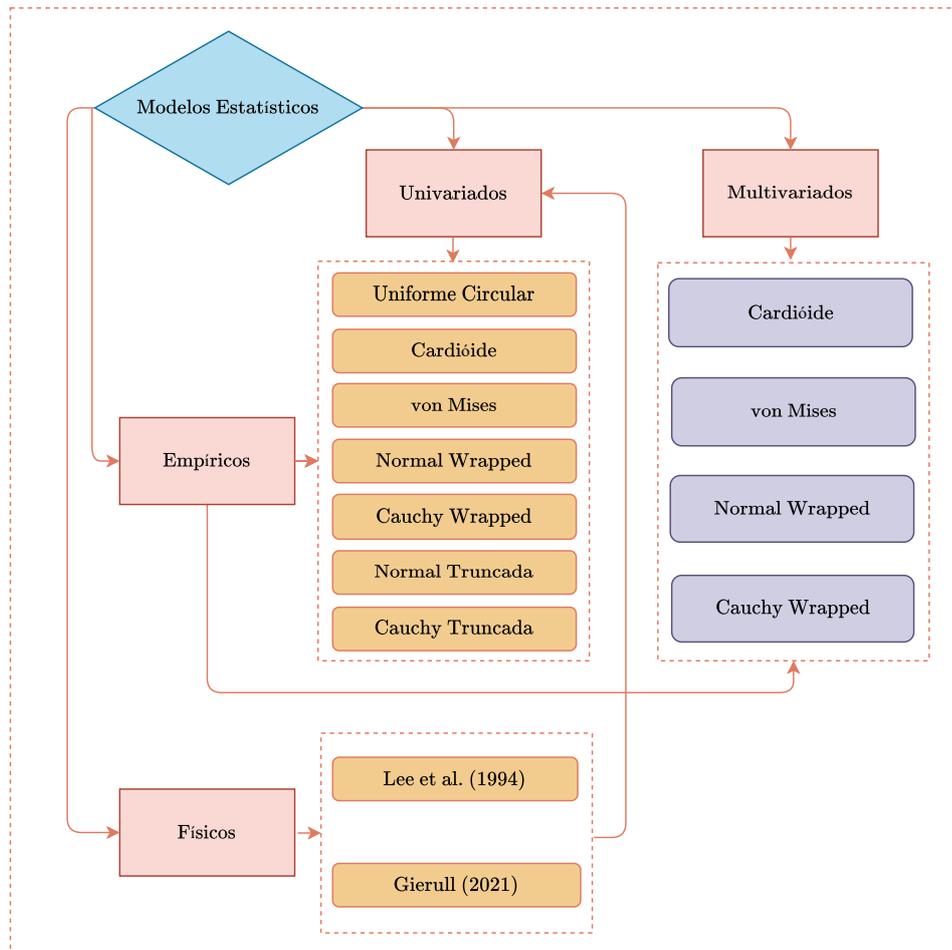
Na Figura 10, além da  $\mathcal{NT}$  e  $\mathcal{CT}$ , apresentamos alguns modelos estatísticos para análise de dados circulares, categorizados em empíricos univariados e multivariados. Outros detalhes sobre os modelos empíricos multivariados, por exemplo, como caracterizar a von Mises com base na Normal Multivariada, a generalização dos modelos, Cardioide e  $\mathcal{CT}$ , podem ser encontrados em: (DOWNS, 1966), (ATEYA; MADHAGI, 2013) e (PAULA et al., 2021).

Na Figura 11, é possível observar as relações entre alguns dos modelos estatísticos apresentados na Figura 10. Por exemplo, os modelos Cardioide, von Mises, Wrapped Normal e Wrapped Cauchy convergem para a Uniforme Circular, como indicado na quarta coluna da Tabela 1. Além disso, com base na Figura 11, podemos notar que, partindo do modelo Normal Padrão com média  $\mu$  e variância  $\sigma^2$ , é possível derivar a  $\mathcal{NT}$ .

Uma das limitações do modelo Cauchy Padrão é a ausência de momentos finitos. Para contornar essa restrição, introduz-se uma versão truncada da distribuição. Como a  $\mathcal{CT}$  é definida em um intervalo finito, ela possui todos os momentos finitos, o que a torna uma alternativa superior ao modelo Cauchy Padrão. Mais detalhes sobre as propriedades dessa distribuição podem ser encontrados em (NADARAJAH; KOTZ, 2006b) e (ATEYA; MADHAGI, 2013).

Os modelos Uniforme Circular, von Mises e Wrapped Normal são amplamente utilizados na modelagem estatística de dados circulares. Uma característica relevante da Uniforme Circular é a ausência de média circular, com o vetor resultante apresentando comprimento médio igual a zero. Já a von Mises destaca-se como uma das escolhas mais comuns para experimentos com

Figura 10 – Modelos estatísticos

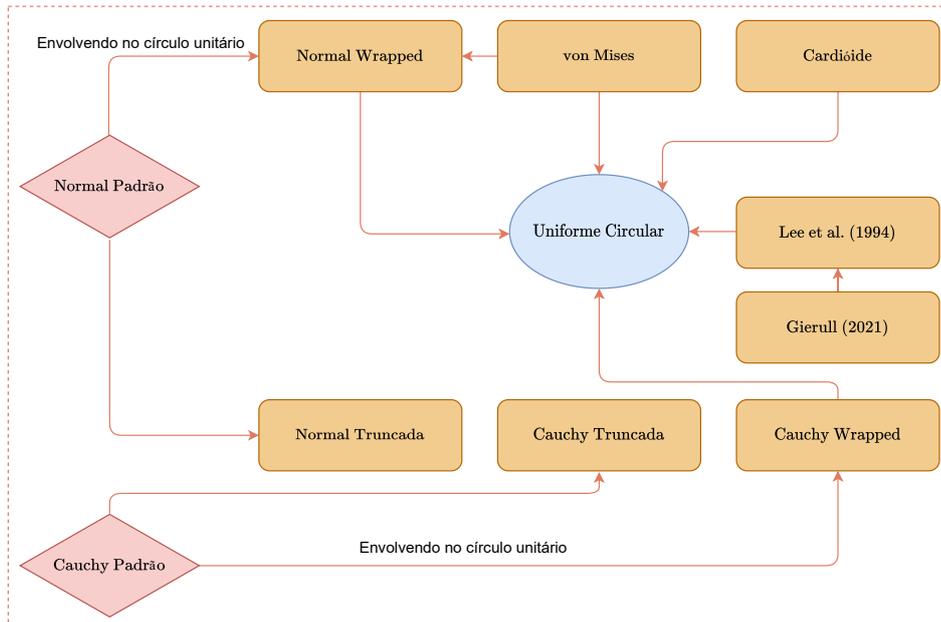


Fonte: O autor (2025).

dados reais, devido às suas boas propriedades, como regularidade, parâmetros interpretáveis e simetria. Mais informações sobre essas propriedades podem ser encontradas em (DOWNS, 1966), (DOWNS; GOULD, 1967), (RAO; SENGUPTA, 2001), (MARDIA; TAYLOR; SUBRAMANIAM, 2006) e (ARTHUR; MARKUS; GRAEME, 2013).

De acordo com as informações da Figura 11 e da Tabela 1, podemos observar que a Wrapped Normal converge para a Uniforme Circular, quando  $\rho \rightarrow 0$ . Além disso, na Tabela 1, apresentamos algumas informações importantes sobre os modelos descritos na Figura 11: a segunda coluna contém os parâmetros dos respectivos modelos, a terceira coluna destaca algumas características relevantes, e a quarta coluna mostra como eles se relacionam.

Figura 11 – Relações entre modelos estatísticos



Fonte: O autor (2025).

Tabela 1 – Características e relações entre modelos estatísticos

Modelo	Parâmetros	Características	Relações
Uniforme Circular	—	Comprimento do vetor sempre igual a zero. Equivalente sob rotação e reflexão.	Ver Figura 11
Cardióide	$-\pi < \mu \leq \pi$ , $-\frac{1}{2} < \rho < \frac{1}{2}$ concentração $\rho$ , média $\mu$	Unimodal e simétrica.	Uniforme para $\rho = 0$
von Mises	$-\pi < \mu \leq \pi$ , $\kappa \geq 0$ concentração $\kappa$ , ângulo médio $\mu$	Depende da função de Bessel. Unimodal e simétrica ao redor de $\mu$ .	Uniforme para $\kappa = 0$ $\mathcal{WN}$ quando $\kappa \rightarrow \infty$
Wrapped Normal	$-\pi < \mu \leq \pi$ , $0 \leq \rho \leq 1$ vector resultante $\rho$ , média $\mu$	Envolvida no círculo unitário. Unimodal e simétrica ao redor $\mu$ .	Uniforme quando $\rho \rightarrow 0$ .
Wrapped Cauchy	$-\pi < \mu \leq \pi$ , $0 \leq \rho \leq 1$ $\rho$ vector resultante, média $\mu$	Envolvida no círculo unitário. Quando $\rho = 1$ concentra-se em $\mu$ .	Uniforme quando $\rho \rightarrow 0$
$\mathcal{NT}$	$-\pi < \mu \leq \pi$ , $\sigma > 0$ $\mu$ locação, $\sigma$ escala	Truncamento unilateral.	Normal Padrão
$\mathcal{CT}$	$-\pi < \mu \leq \pi$ , $\sigma > 0$ $\mu$ locação, $\sigma$ escala	Apresenta todos os momentos finitos.	Cauchy Padrão
(3.20)	$0 <  \rho_c  < 1$ , $-\pi < \psi \leq \pi$ $L$ real positivo	Forma fechada para diferença de fase.	Gamma e GHF
(3.21)	$0 <  \rho_c  < 1$ , $-\pi < \psi \leq \pi$ $L$ inteiro maior do que 1	Idêntica ao modelo (3.20) para $L$ inteiro positivo e maior do que 1.	Gamma e modelo (3.20)

### 3.5 PROCESSO DE FILTRAGEM

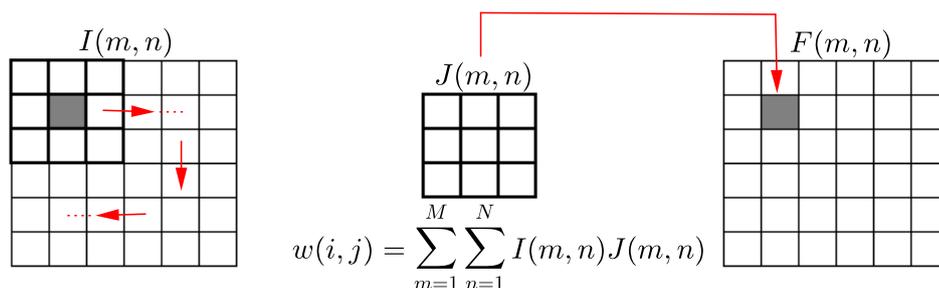
A filtragem em dados SAR busca aprimorar a qualidade das imagens, reduzir ruídos e destacar informações importantes. A filtragem pode ser realizada de diversas maneiras, desde a aplicação de filtros espaciais, temporais e de polarização até o uso de técnicas de processa-

mento de imagens baseadas em estatística. Nas técnicas de filtragem, o contexto é frequentemente considerado de forma local, utilizando janelas móveis para analisar pequenas porções da imagem, ou de forma global, como na transformada de Fourier, que converte a imagem do domínio espacial para o domínio de frequência. Nesta tese, destacamos as técnicas de filtragem no domínio espacial com abordagem local.

Dependendo dos elementos contidos na janela, os filtros podem ser classificados como passa-baixa<sup>2</sup>, que suavizam a imagem ao eliminar detalhes de alta frequência; passa-alta<sup>3</sup>, que realçam bordas e detalhes finos; ou passa-faixa<sup>4</sup>, que preservam apenas uma faixa intermediária de frequências. Para essa análise, são utilizados diversos tipos de janelas móveis, cujo deslocamento é conhecido como operação de convolução ou vizinhança. Nessa operação, cada porção da imagem é processada com base nos pixels vizinhos, permitindo uma análise local e adaptativa.

O processo de filtragem no domínio espacial corresponde a operações de vizinhanças executadas a partir de uma janela que se desloca sobre todos os pixels da imagem, definindo novos valores para cada um desses pixels. O tipo de filtro irá depender essencialmente dos pesos (valores) contidos na janela. Em uma operação de convolução, o tamanho da janela móvel sempre será ímpar, de modo que haja um pixel central, como mostra o exemplo na Figura 12.

Figura 12 – Operação de convolução



Fonte: Adaptado de Zonatta, Ferreira e Zortea (2019)

No exemplo ilustrado na Figura 12, o processo de convolução apresenta o pixel central

- <sup>2</sup> Permitem a passagem de baixas frequências, removendo as altas frequências ou variações inesperadas nos níveis de cinza, o que resulta na suavização da imagem. Exemplos incluem filtros de média, mediana, moda e adaptativos.
- <sup>3</sup> Permitem a passagem de sinais de alta frequência, reduzindo a intensidade de sinais de baixa frequência. Os detectores de bordas são os tipos de filtros mais conhecidos nessa categoria.
- <sup>4</sup> Permitem a passagem de frequências dentro de um intervalo específico, atenuando tanto baixas quanto altas frequências. São utilizados, por exemplo, para realçar padrões texturais ou remover ruídos específicos da imagem.

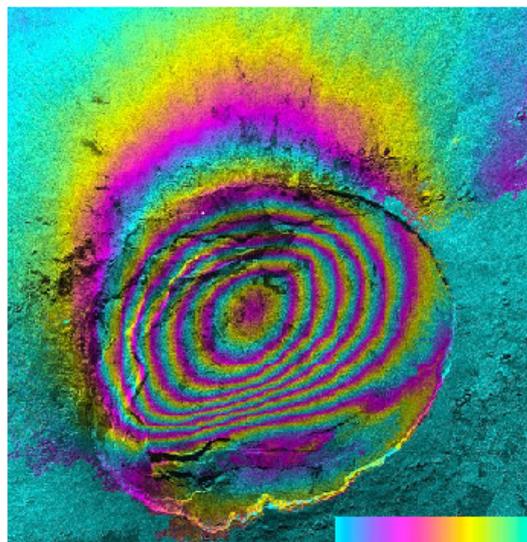
em uma janela  $3 \times 3$  pixels  $w(i, j)$ , destacado em cinza. Os níveis de cinza da imagem  $I(m, n)$  serão multiplicados pelos valores da janela  $J(m, n)$ , os valores dessa multiplicação serão somados e, então, o resultado será inserido na posição correspondente ao pixel central da janela, formando a imagem filtrada  $F(m, n)$  de forma sucessiva.

Pixels localizados nas bordas da cena não possuem todos os vizinhos, permanecendo, dessa forma, inalterados na imagem de saída, ou modificados, mas apenas de acordo com o vizinhos disponíveis, dependendo da abordagem do software (ZONATTA; FERREIRA; ZORTEA, 2019).

### 3.6 FILTROS DE FASE PARA DADOS INSAR

Como definido na subseção 2.4.2, a fase interferométrica, ou interferograma, é uma medida da diferença de fase entre os sinais de radar de duas passagens (ou imagens) do radar sobre uma determinada área imageada. Essa diferença é representada na imagem InSAR como franjas coloridas (Figura 13). Uma franja é um ciclo de cores do espectro, que vai do vermelho ao roxo (vermelho, laranja, amarelo, verde, azul, roxo). Portanto, essas franjas podem ser caracterizadas como ciclos de pixels coloridos ou deformação do solo com linhas de contorno coloridas, as quais, mostram a diferença de fase do sinal recebido pelo instrumento de radar ao longo do tempo.

Figura 13 – Interferograma exibindo ciclos de franjas coloridas



Fonte: Adaptado de (JPL, n.d.)

A filtragem de fase desempenha um papel fundamental em aplicações com dados interferométricos. A precisão na medição de deformações e na determinação da altura do terreno

depende diretamente da qualidade de filtragem. Durante o processo de filtragem do speckle, é essencial preservar boas características importantes da imagem, como as bordas das franjas de fase, que determinam a precisão da altura do terreno (PU et al., 2020).

Conforme dito anteriormente, a diferença de fase entre dois sinais, ou interferograma, contém informação sob a forma de uma fase embrulhada (wrapped phase) módulo  $2\pi$ . O desembrulhamento de fase é então aplicado para remover essa ambiguidade (CHAO; CHEN; LEE, 2013). No entanto, o ruído inerente à aquisição da imagem dificulta o desembrulhamento, especialmente em áreas íngremes. Por esse motivo, a filtragem de fase é considerada uma etapa indispensável e deve ocorrer antes do desembrulhamento da fase.

O desembrulhamento de fase é um importante processo na análise de padrões de franjas, particularmente em técnicas interferométricas. Seu objetivo é converter uma fase embrulhada, caracterizada por descontinuidades, em uma fase contínua. Quando os valores de fase excedem os limites do intervalo  $[-\pi, \pi]$ , ocorre o embrulhamento de fase, resultando em descontinuidades no sinal. O desembrulhamento busca eliminar essas descontinuidades, permitindo uma reconstrução contínua da fase. Para mais informações, consultar (MIGUEL et al., 2002).

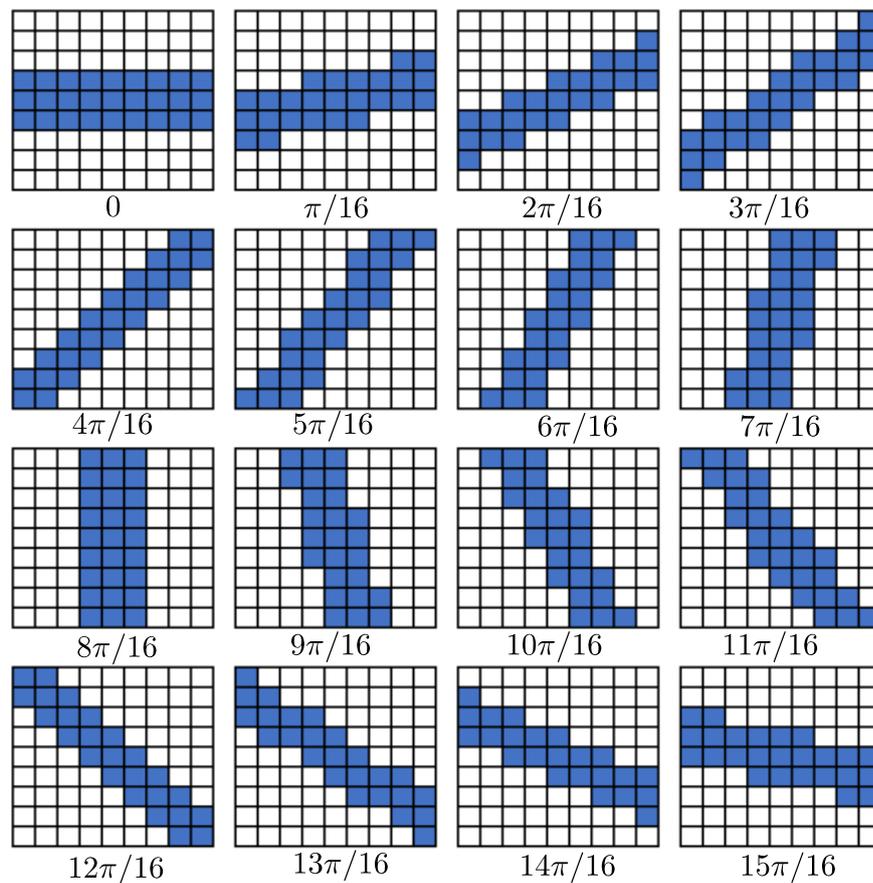
Um filtro para a redução do speckle, deve considerar a manutenção da média da intensidade recebida, a redução da variância e a preservação das bordas e da textura (DOMG; MILNE; FORSTER, 2001). Os métodos de filtragem de fase amplamente utilizados podem ser agrupados em três categorias: métodos de domínio espacial, métodos no domínio de transformação e métodos não locais (NL) (PU et al., 2020). Para mais detalhes sobre filtros não locais consultar (BAIER et al., 2018) e (SUN et al., 2020).

Um método típico de filtragem no domínio espacial é o filtro Lee InSAR, caracterizado por filtrar adaptativamente o ruído ao longo das franjas. Os filtros adaptativos preservam os detalhes (bordas) ao mesmo tempo que reduzem ou suavizam os ruídos. Além disso, nos filtros adaptativos, ao contrário da versão tradicional dos filtros de medianas e médias, o valor do pixel central da janela será substituído apenas se for comprovada a presença de um pixel singular (pixel indesejável, ou seja, pixel afetado pelo ruído) no interior da janela. Para mais informações sobre filtros de fase: Lee InSAR, Goldstein, InSAR-BM3D, Boxcar, consultar, (VELHO; NETO; SIMÕES, 2012), (CHAO; CHEN; LEE, 2013), (PU et al., 2020) e (SUN et al., 2020).

### 3.6.1 Filtro adaptativo Lee InSAR

O filtro Lee InSAR assume o modelo aditivo para o ruído e obedece o critério de minimização local e linear do Erro Quadrático Médio Mínimo (MMSE – Minimum Mean Square Erro) (LEE et al., 1998), (CHAO; CHEN; LEE, 2013). Em contraste com a abordagem da janela quadrada móvel adotada por outros algoritmos, o filtro Lee InSAR utiliza dezesseis janelas, como mostrado na Figura 14. Uma dessas janelas é selecionada, e o pixel central é filtrado com base no critério do MMSE (LEE et al., 1998), (CHAO; CHEN; LEE, 2013).

Figura 14 – Janelas direcionais para o filtro Lee InSAR Refinado



Fonte: Adaptado de (LEE et al., 1998)

O algoritmo associado ao filtro Lee InSAR foi projetado com a finalidade específica de ajustar-se à direção das franjas. Sua composição fundamental abrange as seguintes etapas (CHAO; CHEN; LEE, 2013):

1. Para filtrar o pixel central em uma janela direcional  $9 \times 9$ , as médias do conjunto  $|\langle e^{j\psi} \bullet \rangle|$  são calculadas para todas as 16 janelas da Figura 14 e a janela com a magnitude máxima  $|\langle e^{j\psi} \bullet \rangle|$  é selecionada, fundamentado pelo fato de que, se todas as fases são

idênticas na janela, então  $|\langle e^{j\psi_{\bullet}} \rangle| = 1$ , e se as fases são distribuídas uniformemente, então  $|\langle e^{j\psi_{\bullet}} \rangle| = 0$ , onde  $\psi_{\bullet}$  é o conjunto de valores de pixels em cada janela direcional,  $\langle \cdot \rangle$  representa a média espacial e  $|\cdot|$  denota o módulo de um número complexo.

2. A coerência é obtida a partir do mapa de coerência interferométrica (por exemplo, Figuras (1b), (1d) e (1f) em Lee et al. (1998)). O mapa de coerência quantifica a similaridade entre imagens de radar adquiridas em diferentes momentos ou condições, sendo uma medida derivada da análise dessas imagens, e o desvio padrão do ruído de fase, denotado por  $\sigma_v$ , é selecionado na tabela de pesquisa na Figura 1 (CHAO; CHEN; LEE, 2013) ou Figura 2 (LEE et al., 1994);
3. O MMSE é aplicado a todos os pixels dentro da janela selecionada, e o processo envolve o cálculo de estatísticas locais, como a média e a variância.

### 3.6.2 Filtro Lee InSAR Refinado

O filtro Lee InSAR Refinado (CHAO; CHEN; LEE, 2013) é uma adaptação do filtro Lee InSAR (LEE et al., 1998). O Lee InSAR Refinado inclui um filtro sigma melhorado (para mais detalhes sobre o filtro sigma, consulte (LEE et al., 2009)) e utiliza um limiar para selecionar a janela direcional mais homogênea. O algoritmo é composto pelas seguintes etapas:

#### 1. Seleciona uma janela direcional e um ângulo de janela.

Para filtrar o pixel central em uma janela de tamanho  $\kappa \times \kappa$ , com  $\kappa = 9$ , a janela direcional é selecionada entre 16 opções (Figura 14), com base na magnitude máxima  $|\langle e^{j\psi_{\bullet}} \rangle|$ . Se  $|\langle e^{j\psi_{\bullet}} \rangle|$  é maior que o limiar  $\epsilon_{th}$ , empiricamente definido, então o ângulo da janela é determinado. Caso contrário, o ângulo da janela precisa ser revisado usando uma janela  $\kappa \times \kappa$  com  $\kappa = 11$  e um novo esquema chamado, Ponderação Inversa da Distância (IDW – Inverse Distance Weighting).

Em outras palavras, quando o pixel central não tem um ângulo definido, é necessário recalcular os ângulos de todos os pixels nessa janela. Isso é realizado usando o método de IDW para interpolar o ângulo da janela do pixel não definido com base nos ângulos dos pixels adjacentes ou vizinhos. Para este propósito, as seguintes expressões em (3.28)

são utilizadas:

$$d_{pq} = \frac{1}{d(p, q)}, \quad e^{j\hat{\gamma}} = \frac{\sum_{p,q=1}^{\kappa} d_{pq} e^{j\gamma(p,q)}}{\sum_{p,q=1}^{\kappa} d_{pq}}, \quad (3.28)$$

em que  $d(p, q)$  é distância Euclidiana do pixel central  $(m, n)$  até pixels arbitrários  $(p, q)$  na mesma janela, e  $\gamma(p, q)$  representa o ângulo da janela de um pixel de coordenadas  $(p, q)$  dentro da janela  $\kappa \times \kappa$ . Portanto, uma das 16 janelas cujo ângulo é mais próximo do ângulo ponderado  $\hat{\gamma}$  será designada como a nova janela. Por exemplo, a janela 3, cujo ângulo é igual a  $3\pi/16$  (Figura 14), é selecionada se o ângulo  $\hat{\gamma}$  for igual a  $2.9\pi/16$ .

## 2. Seleciona pixels adequados na janela direcional.

A ideia foi extraída do filtro sigma melhorado desenvolvido para reduzir o efeito do speckle em imagens SAR (LEE et al., 2009). Primeiramente, para um intervalo sigma, somente pixels dentro do intervalo  $(\psi_c - \psi_\xi, \psi_c + \psi_\xi]$  serão incluídos no processamento, onde  $\psi_c$  é a fase associada ao pixel central em uma janela direcional. Em segundo lugar,  $\psi_\xi$  é calculado a partir do modelo em (3.20) para um dado valor de  $\xi$ .

No algoritmo Lee InSAR (subseção 3.6.1), todos os pixels da janela selecionada são considerados na estimativa estatística. No entanto, segundo Chao, Chen e Lee (2013), resultados insatisfatórios podem ocorrer se  $\psi_c$  for um pixel singular. Para resolver esse problema, o algoritmo Lee Refinado aplica um filtro médio  $3 \times 3$  para ajustar o pixel central. Após desembrulhar a fase dentro da janela  $3 \times 3$ , os nove pixels são ordenados como  $\psi_{(1)} \leq \psi_{(2)} \leq \psi_{(3)} \leq \dots \leq \psi_{(7)} \leq \psi_{(8)} \leq \psi_{(9)}$ . Se  $\psi_c$ , estiver localizada fora da faixa entre  $\psi_{(3)}$  e  $\psi_{(7)}$ , ou seja,  $\psi_c < \psi_{(3)}$  ou  $\psi_c > \psi_{(7)}$ , então  $\psi_c$  é substituída pela média dos valores  $\psi_{(3)}, \psi_{(4)}, \psi_{(5)}, \psi_{(6)}$ , e  $\psi_{(7)}$ . Caso contrário, o pixel central permanece inalterado.

## 3. Aplicação do MMSE

O MMSE é aplicado aos pixels selecionados no intervalo sigma  $(\psi_c - \psi_\xi, \psi_c + \psi_\xi]$ . Em seguida, os valores locais, da média circular amostral e da variância são calculados. Para mais informações sobre o filtro Lee InSAR Refinado, consultar (CHAO; CHEN; LEE, 2013) e (LI; CHEN; LEE, 2020).

### 3.6.3 Algoritmos de Filtragem Propostos

O filtro InSAR Refinado Aprimorado (LI; CHEN; LEE, 2020) é uma melhoria do filtro InSAR Refinado (CHAO; CHEN; LEE, 2013). Ele consiste em duas etapas: (i) selecionar um subconjunto de observações espaciais contíguas e homogêneas, e (ii) calcular a média circular truncada dessas observações. Nossa proposta entra na segunda etapa, portanto, omitiremos os detalhes da seleção da amostra.

Suponha que tenhamos selecionado a amostra  $\psi' = (\psi'_1, \psi'_2, \dots, \psi'_N)$  das fases observadas. Para a filtragem, as fases  $\psi_z$  são selecionadas dentro do intervalo  $\psi'_z = (\psi_c - \psi_\xi, \psi_c + \psi_\xi]$ , onde  $\psi_\xi$  é o limite de fase correspondente à proporção desejada  $\xi$ . O valor de  $\psi_\xi$  é calculado numericamente resolvendo:

$$F(\psi_\xi) - F(-\psi_\xi) = \xi, \quad (3.29)$$

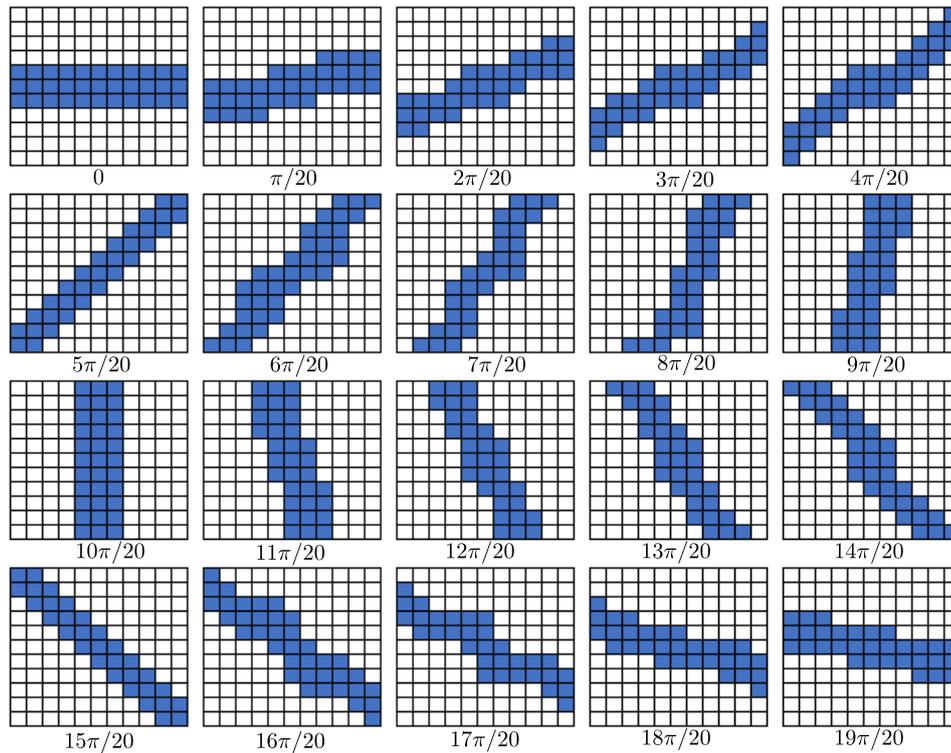
em que  $F$  é a CDF do modelo adotado. Segundo Chao, Chen e Lee (2013), o valor de  $\psi_\xi$  pode ser obtido numericamente a partir de  $\xi$  para todos os pixels. Em outras palavras, o limite de fase pode ser calculado uma única vez para toda a imagem. De acordo com Li, Chen e Lee (2020), o valor ótimo para  $\xi$  é 0.9.

Com base no nível de ruído local, este estudo apresenta três métodos para a filtragem do ruído de fase: TcNfilter, TcCfilter e LInSARRFE. Os dois primeiros utilizaram os modelos empíricos  $\mathcal{NT}$  e  $\mathcal{CT}$  para calcular o limite de fase, respectivamente. Enquanto, o LInSARRFE calcula  $\psi_\xi$  por meio da expressão definida em (3.21) para valores inteiros de  $L$  maiores do que 1. Para os filtros propostos, foram criadas 20 janelas direcionais  $11 \times 11$  (Figura 15), e um algoritmo foi desenvolvido para selecionar aquela que contém pixels com fases mais homogêneas (LI; CHEN; LEE, 2020).

Análogo ao Lee Refinado, a escolha da janela direcional é baseada na magnitude máxima  $|\langle e^{j\psi} \rangle|$ . Se a magnitude da média do exponencial complexo  $|\langle e^{j\psi} \rangle|$  exceder o limiar  $\epsilon_{th}$ , empiricamente definido, o ângulo de orientação da janela,  $\phi_w = n\pi/20, n = 0, 1, \dots, 19$ , é então determinado. Caso contrário, o ângulo da janela precisa ser revisado usando uma janela  $\kappa \times \kappa$ , com  $\kappa = 13$ , para fornecer uma estimativa mais confiável da direção das franjas, usando o IDW. Para este propósito, utilizamos as expressões em (3.28).

Para manter a resolução, propomos aplicar o estimador MMSE (LEE et al., 2009) aos pixels selecionados em  $\psi'_z$  com base no limite de fase. A equação (3.30) é aplicada para obter a fase

Figura 15 – Janelas direcionais para os filtros propostos



Fonte: Adaptado de (LI; CHEN; LEE, 2020)

filtrada.

$$\hat{\psi}_x = \bar{\psi}'_z + b(\psi_z - \bar{\psi}'_z), \quad (3.30)$$

em que,

$$b = \frac{\text{var}(\psi_x)}{\text{var}(\psi'_z)} \quad \text{e} \quad \text{var}(\psi_x) = \text{var}(\psi'_z) - \sigma_v^2 \quad (3.31)$$

sendo,  $b \in [0, 1]$  o peso, com  $\bar{\psi}'_z$  e  $\text{var}(\psi'_z)$  representando, respectivamente, a média circular local amostral e a variância (I., 1995), Apêndice B), enquanto  $\bar{\psi}_v$  e  $\sigma_v^2$  denotam a média circular e a variância do ruído:

$$\sigma_v^2 = \int_{-\psi_\xi}^{\psi_\xi} (\psi - \bar{\psi}_v)^2 f(\psi) d\psi, \quad \text{e} \quad \bar{\psi}_v = \int_{-\psi_\xi}^{\psi_\xi} \psi f(\psi) d\psi. \quad (3.32)$$

Na página 149 do livro *Polarimetric Radar Imaging*, Lee e Pottier (2009) apresentam a dedução da equação (3.30). Conforme definido em (3.31), o parâmetro  $b^5$  depende diretamente da relação entre a variância local e a variância do ruído. De acordo com Lee e Pottier (2009), quando  $b = 1$ , o filtro preserva a valor original da fase (pixel), minimizando o efeito da filtragem e  $\sigma_v^2 = 0$ , por outro lado, quando  $b = 0$ , a fase filtrada é substituída pela média local e  $\text{var}(\psi'_z) = \sigma_v^2$ .

<sup>5</sup> Fator adaptativo que controla a intensidade da filtragem com base na variância da fase interferométrica.

Para assegurar um comportamento adequado do filtro, o peso  $b$  deve pertencer ao intervalo  $[0, 1]$ . Caso contrário, a filtragem pode ter efeitos indesejados. Se  $b < 0$ , o ruído pode ser ampliado em vez de reduzido, resultando em descontinuidades. Por outro lado, se  $b > 1$ , a filtragem superestima a média local, causando um efeito de borramento ou suavização excessiva na imagem. Além disso, a equação (3.30) depende da variância  $\text{var}(\psi_x)$ , que pode se tornar negativa devido a um número insuficiente de amostras ou ao uso de um valor excessivamente alto para  $\sigma_v^2$ . Quando isso ocorre,  $\text{var}(\psi_x)$  é ajustada para zero, garantindo que  $b$  permaneça dentro do intervalo adequado (LEE; POTTIER, 2009).

O Algoritmo 1 apresenta de maneira detalhada os filtros propostos, proporcionando uma visão abrangente das etapas e processos envolvidos em sua implementação.

---

**Algoritmo 1:** Métodos Propostos: Etapas

---

**Entrada:** Imagem interferométrica, modelo padrão  $F: \mathcal{NT}, \mathcal{CT}$ , ou (3.21), proporção  $0 < \xi \leq 1$ , limiar  $\epsilon_{th}$  definido empiricamente

**Saída:** Imagem filtrada

**Computar:** Filtrar cada pixel central  $\psi_c$  usando uma janela  $11 \times 11$  pixels (Figura 15 (LI; CHEN; LEE, 2020))

- 1 Calcular o limite de fase  $\psi_\xi$  tal que  $F(\psi_\xi) - F(-\psi_\xi) = \xi$
- 2 **for** cada janela direcional **do**
- 3     Calcular  $|\langle e^{i\psi_\bullet} \rangle|$ , onde  $\psi_\bullet$  denota o conjunto de valores de pixels em cada janela direcional
- 4     **if**  $|\langle e^{i\psi_\bullet} \rangle| < \epsilon_{th}$  **then**
- 5         Recalcular o ângulo com IDW (Equações em (3.28)) usando uma janela de  $13 \times 13$
- 6         Selecionar o ângulo da janela mais próximo de  $\hat{\gamma}$
- 7     Dividir em subjanela de  $3 \times 3$ , desembrulhar a fase, ordenar os pixels
- 8     **if**  $\psi_c < \psi_{(3)}$  **ou**  $\psi_c > \psi_{(7)}$  **then**
- 9         O pixel central  $\psi_c$  é singular; portanto, substituí-lo pela média dos pixels no intervalo  $[\psi_{(3)}, \psi_{(7)}]$
- 10     Selecionar pixels dentro da janela direcional conforme o intervalo  $\psi_c - \psi_\xi < \psi_z \leq \psi_c + \psi_\xi$
- 11     Calcular estatísticas locais:  $\overline{\psi'_z}$  e  $\text{var}(\psi'_z)$
- 12     Calcular média e variância do ruído:  $\overline{\psi}_v$  e  $\sigma_v^2$
- 13     Calcular peso  $b$  e atualizar  $\hat{\psi}_x$  usando o estimador MMSE

**Saída:** Retornar a imagem filtrada

---

### 3.7 DIVERGÊNCIA ESTATÍSTICA

Nesta seção, exploramos distâncias estatísticas para quantificar a discrepância entre modelos probabilísticos. Quando essas medidas incluem a propriedade de simetria, são referidas como distâncias estocásticas. Estas são capazes de medir a distância entre dois modelos de probabilidade, assim, quanto maior a discrepância entre os dois modelos, maior será a distância entre eles.

No estudo realizado por Frery, Nascimento e Cintra (2014), foram investigadas distâncias estocásticas baseadas na classe  $(h, \phi)$  de divergências para lidar com problemas de testes de hipóteses em dados contaminados com ruído speckle. Especificamente, foram consideradas as distâncias de Kullback-Leibler (KL), Rényi, Jensen-Shannon, Hellinger, Bhattacharyya, Aritmética-Geométrica, Triangular e Média Harmônica. Neste estudo, estamos interessados em mostrar e avaliar o melhor e o pior ajuste (em termos de distância de KL) entre a expressão (3.20) em comparação com o modelo em (3.21).

Segundo Frery, Nascimento e Cintra (2014), a expressão geral para a classe de divergências  $(h, \phi)$  entre as densidades  $f_X$  e  $f_Y$  pode ser definida de acordo com a equação,

$$D_{\phi}^h(\mathbf{X}, \mathbf{Y}) = h \left( \int_{\mathcal{A}} \phi \left( \frac{f_X(Z; \boldsymbol{\theta}_1)}{f_Y(Z; \boldsymbol{\theta}_2)} \right) f_Y(Z; \boldsymbol{\theta}_2) dZ \right) \quad (3.33)$$

em que  $\mathbf{X}$  e  $\mathbf{Y}$  são matrizes aleatórias definidas sobre o mesmo espaço probabilístico associadas às densidades  $f_X$  e  $f_Y$ , respectivamente, com mesmo suporte  $\mathcal{A} \subset \mathbb{R}$ . A função  $h : (0, \infty) \rightarrow [0, \infty)$  é convexa, estritamente crescente e  $h(0) = 0$ .

#### 3.7.1 Distâncias Estocásticas

Escolhendo adequadamente as funções  $h$  e  $\phi$  na equação em (3.33), é possível obter diversas divergências. No entanto, é importante observar que essas métricas nem sempre são medidas de distância, uma vez que a propriedade de simetria e a desigualdade triangular nem sempre são cumpridas. Apesar disso, essas medidas são eficientes para comparar diferentes modelos de probabilidade. De acordo com Frery, Nascimento e Cintra (2014), as divergências resultantes em (3.33) podem ser simetrizadas de modo direto utilizando a equação dada por:

$$d_{\phi}^h(\mathbf{X}, \mathbf{Y}) = \frac{D_{\phi}^h(\mathbf{X}, \mathbf{Y}) + D_{\phi}^h(\mathbf{Y}, \mathbf{X})}{2}. \quad (3.34)$$

As funções  $d_{\phi}^h : \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{R}$  são distâncias sobre  $\mathcal{A}$  pois, para todo  $\mathbf{X}, \mathbf{Y} \in \mathcal{A}$ , as seguintes propriedades são válidas:

1.  $d_{\phi}^h(\mathbf{X}, \mathbf{Y}) \geq 0$  (não-negatividade);
2.  $d_{\phi}^h(\mathbf{X}, \mathbf{Y}) = d_{\phi}^h(\mathbf{Y}, \mathbf{X})$  (simetria);
3.  $d_{\phi}^h(\mathbf{X}, \mathbf{Y}) = 0 \Leftrightarrow \mathbf{X} = \mathbf{Y}$  (definitividade).

Como mencionado anteriormente, optamos por utilizar a medida de divergência de KL. A expressão geral para a distância estocástica empregada neste trabalho é apresentada em (3.35), conforme discutido por Frery, Nascimento e Cintra (2014),

$$d_{\text{KL}}(\mathbf{X}, \mathbf{Y}) = \frac{1}{2} \int (f_{\mathbf{X}} - f_{\mathbf{Y}}) \log \left( \frac{f_{\mathbf{X}}}{f_{\mathbf{Y}}} \right). \quad (3.35)$$

### 3.8 PROCEDIMENTOS E MÉTODOS

Os procedimentos adotados nesta tese foram organizados em três etapas principais: (i) levantamento e modelagem estatística, (ii) aplicação dos modelos e filtros propostos e (iii) avaliação do desempenho das abordagens utilizadas. A seguir, detalhamos cada uma dessas etapas.

#### 3.8.1 Levantamento e Modelagem Estatística

Inicialmente, realizamos uma revisão da literatura com o objetivo de identificar e mapear modelos estatísticos disponíveis para dados circulares. Em seguida, os parâmetros dos modelos físicos, apresentados em (3.20) e (3.21), foram estimados utilizando a técnica de MLE. Para os modelos empíricos  $\mathcal{NT}$  e  $\mathcal{CT}$ , fixamos o parâmetro de locação em  $\mu = 0$ , enquanto o parâmetro de escala  $\sigma$  foi estimado via MLE. Além disso, para mensurar a diferença entre os modelos, utilizamos a distância de Kullback-Leibler e o erro absoluto.

#### 3.8.2 Aplicação dos Modelos e Filtros

Os filtros propostos, LInSARRFE, TcNfilter e TcCfilter, juntamente com Lee Refinado, foram aplicados tanto em dados sintéticos quanto em dados SAR. Para a aplicação dos filtros fundamentados em modelos físicos, o primeiro passo consistiu a geração do mapa de coerência a partir dos interferogramas em estudos. Esse mapa foi obtido com base nas estimativas de  $|\rho_c|$ , que representam a coerência entre os sinais, sendo calculadas para cada pixel utilizando o

método de máxima verossimilhança (MLE). Este procedimento foi realizado para as três aplicações analisadas. Para os filtros baseados nos modelos empíricos ( $\mathcal{NT}$  e  $\mathcal{CT}$ ), os parâmetros foram estimados conforme descrito na etapa anterior.

Após a aplicação dos filtros, para a restauração das informações de fase, aplicamos o processo de desembrulhamento (MIGUEL et al., 2002). Esse procedimento é essencial na análise de dados interferométricos, utilizado para corrigir ambiguidades decorrentes da periodicidade natural das medições de fase obtidas por radar, restritas ao intervalo  $(-\pi, \pi]$ , gerando descontinuidades módulo  $2\pi$ .

Para o desembrulhamento local da fase, adotamos um método simplificado, adaptado de Miguel et al. (2002), na etapa 7 do Algoritmo 1. Esse método remove progressivamente os saltos de  $2\pi$  ao longo da matriz, garantindo a continuidade da fase. Nesse método, a fase desembrulhada  $\psi_u(i, j)$  é obtida a partir da fase embrulhada  $\psi(i, j)$  usando um método sequencial que percorre a matriz linha por linha (da esquerda para a direita) e, quando necessário, na direção vertical (de cima para baixo). Matematicamente, temos:

$$\psi_u(i, j) = \begin{cases} \psi(i, j), & \text{se } i = 1 \text{ e } j = 1 \text{ (primeiro pixel)} \\ \psi_u(i, j - 1) + \Delta_x(i, j), & \text{se } j > 1 \text{ (prioridade na direção horizontal)} \\ \psi_u(i - 1, j) + \Delta_y(i, j), & \text{se } j = 1 \text{ e } i > 1 \text{ (primeira coluna, direção vertical)} \end{cases}$$

onde os ajustes de fase são definidos por:

$$\Delta_x(i, j) = (\psi(i, j) - \psi(i, j - 1)) + \begin{cases} -2\pi, & \text{se } (\psi(i, j) - \psi(i, j - 1)) > \pi \\ +2\pi, & \text{se } (\psi(i, j) - \psi(i, j - 1)) < -\pi \\ 0, & \text{caso contrário} \end{cases}$$

$$\Delta_y(i, j) = (\psi(i, j) - \psi(i - 1, j)) + \begin{cases} -2\pi, & \text{se } (\psi(i, j) - \psi(i - 1, j)) > \pi \\ +2\pi, & \text{se } (\psi(i, j) - \psi(i - 1, j)) < -\pi \\ 0, & \text{caso contrário} \end{cases}$$

### 3.8.3 Avaliação do Desempenho

Para avaliar o desempenho dos métodos propostos, utilizamos diferentes métricas para dados simulados e SAR:

- **Dados Simulados:** Comparação com uma imagem de referência (livre de ruídos) utilizando as métricas:

1. Índice de Similaridade Estrutural (SSIM): Avalia a similaridade entre duas imagens considerando luminância, contraste e estrutura. Seu valor varia entre  $[-1, 1]$ , onde 1 indica que imagens são idênticas. De acordo com Vitale et al. (2022), o SSIM é definido como:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3.36)$$

em que,

- $\mu_x$  e  $\mu_y$  são as médias das imagens  $x$  e  $y$ ;
- $\sigma_x^2$  e  $\sigma_y^2$  são as variâncias das imagens  $x$  e  $y$ ;
- $\sigma_{xy}$  é a covariância entre as imagens  $x$  e  $y$ ;
- $C_1$  e  $C_2$  são constantes de estabilização numérica. Quanto maior a faixa dinâmica (por exemplo, com 8 bits,  $R = 255$ ), maiores serão os valores de  $C_1$  e  $C_2$ , reduzindo a sensibilidade a pequenas variações. As constantes são definidas como:

$$C_1 = (k_1R)^2, \quad C_2 = (k_2R)^2$$

onde,

- \*  $R$  é a faixa dinâmica dos valores de pixel ( $R = 2^{\text{\#bits por pixel}} - 1$ );
  - \*  $k_1 = 0.01$  e  $k_2 = 0.03$  são as constantes padrão para evitar instabilidades numéricas.
2. Raiz do Erro Quadrático Médio (RMSE): Mede a diferença absoluta entre os valores dos pixels das imagens comparadas. Quanto menor o RMSE, mais próxima a imagem filtrada está da de referência (livre de ruídos):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \quad (3.37)$$

3. Média da Diferença (MD): Analisa o viés médio entre as imagens comparadas, identificando possíveis tendências de super ou subestimação da fase.

$$MD = \frac{1}{N} \sum_{i=1}^N (x_i - y_i) \quad (3.38)$$

onde, quanto menor a média, mais próximos os valores filtrados estão dos originais.

4. Variância da Diferença (VD): Mede a dispersão dos erros entre as imagens, fornecendo informações sobre a estabilidade do erro ao longo da imagem.

$$VD = \frac{1}{N} \sum_{i=1}^N [(x_i - y_i) - MD]^2 \quad (3.39)$$

onde, uma variância baixa indica maior consistência nos dados:

5. Tempo de Execução (RT): Representa o tempo necessário para a execução do processo de filtragem, avaliando a eficiência computacional dos filtros aplicados.

Essas métricas permitem uma avaliação quantitativa da filtragem, facilitando a comparação entre diferentes métodos quanto à precisão, eficiência e custo computacional.

- **Dados SAR:** Avaliação baseada no número de resíduos (inconsistência de fase) presentes na imagem filtrada e no tempo de execução (RT). Segundo Li, Chen e Lee (2020), o número de resíduos na imagem filtrada é sempre um bom critério para avaliar a eficiência dos algoritmos de filtragem, especialmente em cenários onde a fase limpa (referência) não é conhecida.

Os resíduos podem comprometer efetivamente o desembrulhamento de fase. Portanto, um bom filtro deve reduzir seu número, facilitando a reconstrução da informação de fase. Quando a imagem de referência não está disponível, o número de resíduos é utilizado para avaliar a qualidade do processo de filtragem em dados SAR (VITALE et al., 2022).

O número de resíduos é calculado a partir do rotacional (soma das diferenças de fase entre pixels adjacentes dentro de uma janela) da fase envolvida em uma janela de  $2 \times 2$ . Primeiramente, as diferenças de fase entre dois pixels adjacentes são calculadas em  $(-\pi, \pi]$ . Se a diferença entre dois pixels ultrapassar os limites do intervalo  $(-\pi, \pi]$ , a fase será ajustada automaticamente para garantir que as diferenças de fase fiquem dentro dos limites estabelecidos.

Em imagens interferométricas, quando o rotacional é diferente de zero, indica a presença de resíduos (ou erros) na fase. Isso acontece quando há uma descontinuidade na fase ao redor da janela analisada. O número de resíduos é obtido contando quantas janelas  $2 \times 2$  apresentam um rotacional diferente de zero. Quanto maior o número de resíduos, maior a quantidade de descontinuidade ou erro na fase.

## 4 RESULTADOS

Neste capítulo, utilizamos duas imagens SAR fornecidas pelo Sentinel-1, banda C, e uma imagem simulada para testar as abordagens propostas e compará-las com o filtro Lee Refinado. Primeiramente, analisamos os dados simulados (VITALE et al., 2022). Em seguida, analisamos os dados da erupção do vulcão La Cumbre, localizado na Ilha Fernandina, arquipélago de Galápagos, Equador, ocorrido em 12 de janeiro de 2020 (PECHNIKOV, 2020). Por último, a região estudada, abrange a área afetada por um terremoto ocorrido na fronteira Irã-Iraque, com a cidade curda iraquiana de Halabja, em 12 de novembro de 2017 (PECHNIKOV, 2017).

### 4.1 DISTÂNCIA ENTRE MODELOS

Para mensurar a diferença entre os modelos físicos definidos em (3.20) e (3.21), empregamos a distância de Kullback-Leibler ( $d_{KL}$ ). A  $d_{KL}$  mede a diferença entre dois modelos, avaliando o grau de divergência entre eles. Verificar essa discrepância permite entender até que ponto a expressão em (3.21), representa com fidedignidade o modelo em (3.20). A Tabela 2 apresenta uma comparação entre o modelo definido em (3.20) e a expressão com propriedades mais tratáveis dada em (3.21) (GIERULL, 2021), assumindo valores aproximados para o número de looks, definidos empiricamente.

#### 4.1.1 Distância entre os modelos (3.20) e (3.21)

Para calcular a distância entre os modelos físicos, adotamos os parâmetros fixos  $|\rho_c| = 0,7$ ,  $\theta = 0$ , variando apenas o valor de  $L$ , conforme Tabela 2. Os dados apresentados na Tabela 2 e ilustrados na Figura 16 mostram que, à medida que o número de looks aumenta, a  $d_{KL}$  diminuem. Isso indica que as expressões (3.20) e (3.21) se aproximam, sugerindo que ambas convergem para valores semelhantes conforme o número de looks aumenta.

Entretanto, a última linha da Tabela 2 destaca um ponto importante: para um número de looks superior a 171, ocorre um erro ao integrar a expressão definida em (3.20). Isso significa que não é possível calcular a distância entre as duas expressões, indicando uma instabilidade numérica e limitando a aplicabilidade do modelo dado por Lee et al. (1994) para um número grande de looks. Essa instabilidade está relacionada à presença da GHF, cujos cálculos se

tornam complexos para  $L$  grande ou valores de correlação elevados.

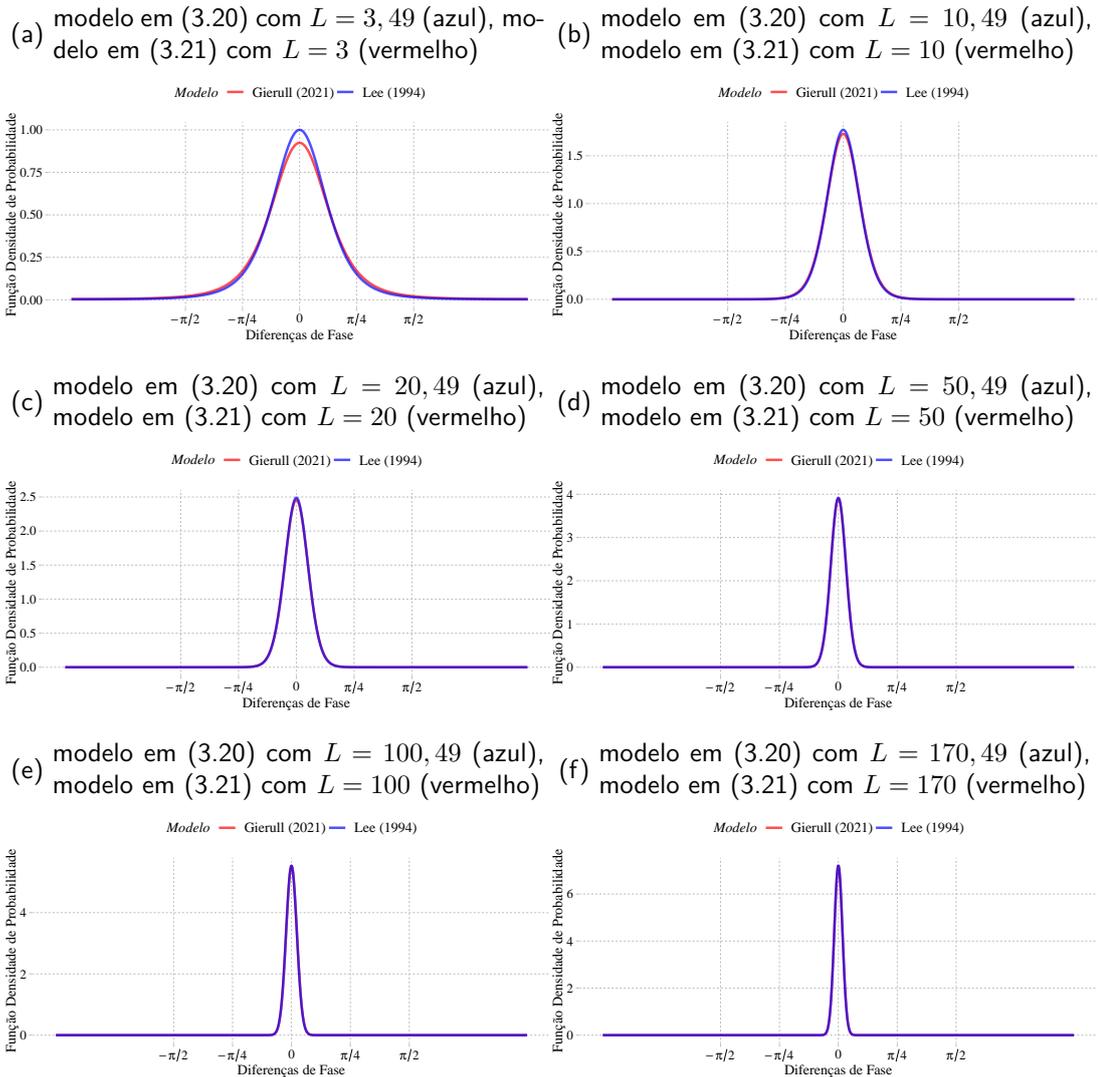
Tabela 2 – Distância entre modelos físicos truncando o número de looks

Looks $L$		Distância de Kullback-Leibler $d_{KL}$
Eq. (3.21)	Eq. (3.20)	
3	3,49	0,0054
10	10,49	0,0006
20	20,49	0,0002
50	50,49	$2,4 \times 10^{-5}$
100	100,49	$5,8 \times 10^{-6}$
170	170,49	$2,1 \times 10^{-6}$

Para  $L > 171$ , ocorre um erro no modelo em (3.20).

Com base nos dados apresentado na Tabela 2, a distância  $d_{KL}$  entre os modelos definidos em (3.20) e (3.21) revela aspectos importantes, descritos a seguir. Observamos que, à medida que o número de looks aumenta, a  $d_{KL}$  diminui gradativamente, apontando uma maior proximidade entre os modelos. Por exemplo, para  $L = 3$ , a  $d_{KL}$  é 0,0054, enquanto para  $L = 170$ , é  $2,1 \times 10^{-6}$ , sugerindo que, para altos valores de  $L$ , as expressões tendem a convergir, ou seja, a perda em utilizar o modelo (3.21) em comparação com o modelo (3.20) é praticamente irrelevante. Os dados apresentados na Tabela 2 e ilustrados na Figura 16 mostram a relação entre a  $d_{KL}$  e o número de looks, evidenciando que, à medida que  $L$  aumenta, os modelos se tornam cada vez mais similares.

Figura 16 – Ajustes entre os modelos físicos

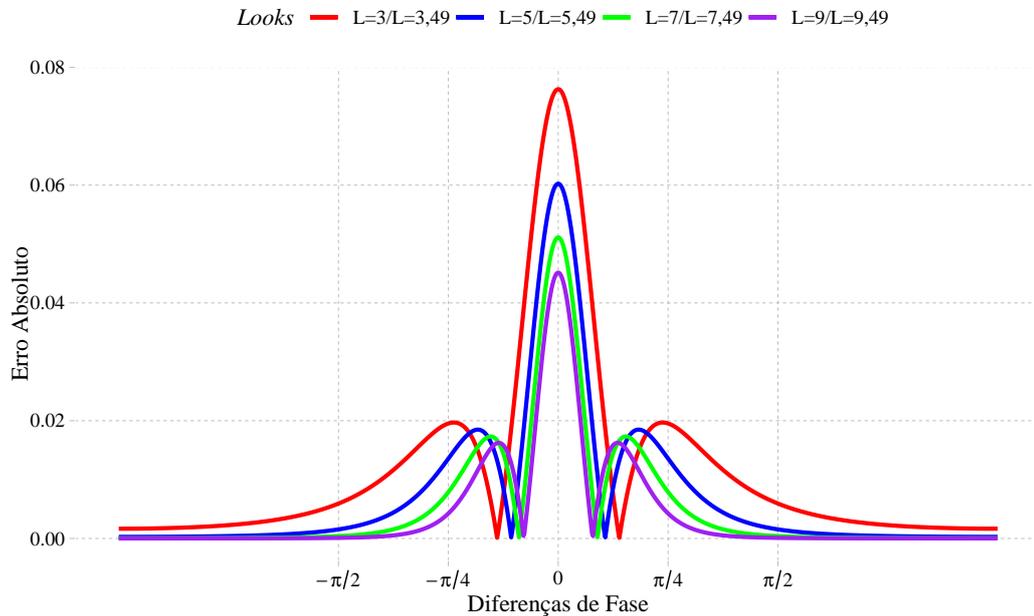


Adicionalmente, a Figura 17 mostra os erros absolutos ponto a ponto entre a expressão em (3.20), indexada por  $L = 3, 49; 5, 49; 7, 49; 9, 49$ , e a expressão em (3.21), assumindo  $L = 3; 5; 7; 9$ . Como podemos observar na Figura 17, a diminuição do erro à medida que o valor de  $L$  aumenta sugere que, para valores inteiros de  $L$  maiores que 1, a aplicação do modelo em (3.21) não resulta em uma perda significativa de precisão em comparação com o modelo em (3.20), de modo que ao aumentarmos o número de looks, a precisão do modelo em (3.21) se aproxima ou converge em relação ao modelo em (3.20), resultando em diferenças absolutas menores entre os modelos.

O erro absoluto reflete a discrepância numérica entre os resultados previstos pelos modelos. De acordo com a Figura 17, o maior erro ocorre quando  $L = 3$  para o modelo em (3.21) e  $L = 3, 49$  para o modelo em (3.20) (vermelho), enquanto o menor erro é alcançado quando

$L = 9$  para o modelo em (3.21) e  $L = 9,49$  para o modelo em (3.20) (roxo).

Figura 17 – Erro Absoluto entre os modelos em (3.20) e (3.21):  $L = 3/L = 3,49$  (vermelho),  $L = 5/L = 5,49$  (azul),  $L = 7/L = 7,49$  (verde) e  $L = 9/L = 9,49$  (roxo)



Fonte: O autor (2025).

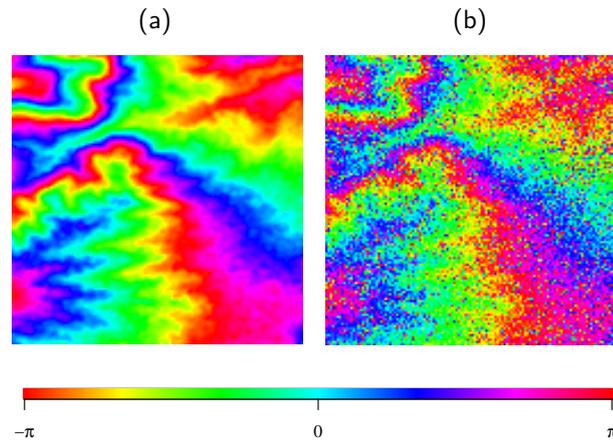
## 4.2 FILTRAGEM DE DADOS SIMULADOS

Para avaliar a eficiência dos métodos propostos, utilizamos uma imagem simulada com dimensões de  $128 \times 128$  pixels. A imagem mostrada em 18(a), livre de ruído, serviu como referência. Por outro lado, à imagem apresentada em 18(b) foi adicionado ruído, simulado pelo modelo em (3.21), de forma uniforme em toda a sua extensão, com a coerência fixada em 0.6. Para gerar valores pseudoaleatórios do modelo em (3.21) utilizamos a técnica de aceitação-rejeição <sup>1</sup>.

As figuras a seguir ilustram os resultados obtidos após a aplicação dos filtros avaliados. As Figuras 19(a), 19(b), 19(c) e 19(d) apresentam a imagem filtrada correspondente à Figura 18(b), utilizando, respectivamente, os filtros Lee InSAR Refinado, LInSARRFE, TcNfilter e TcCfilter. Além disso, as Figuras 19(e), 19(f), 19(g) e 19(h) exibem as fases desembrulhadas geradas a partir das imagens filtradas.

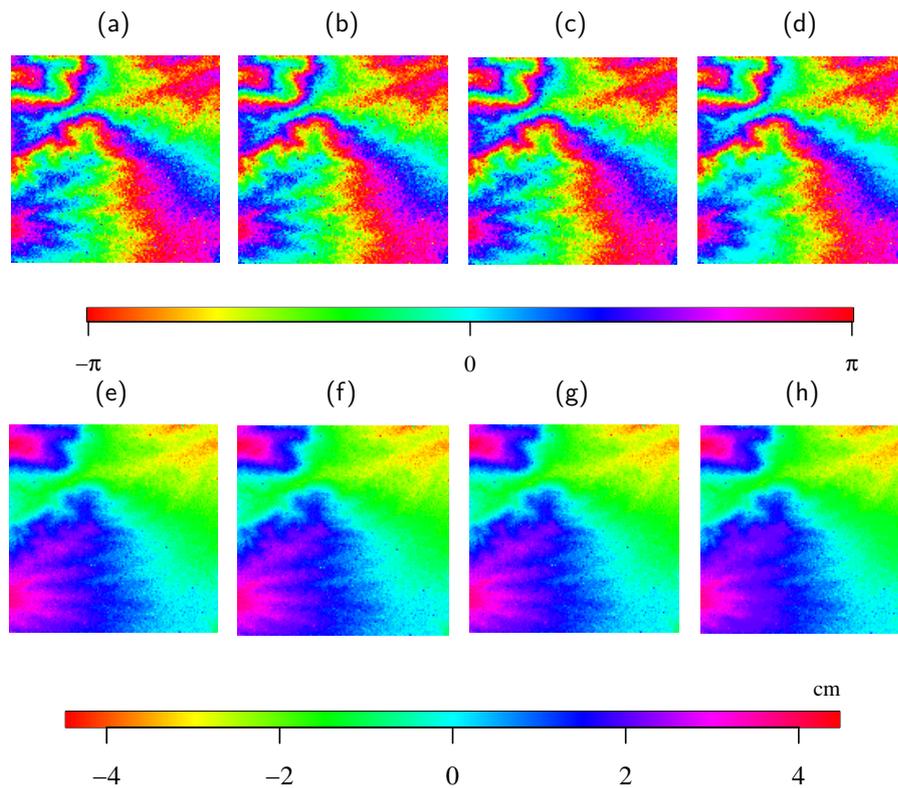
<sup>1</sup> A técnica de aceitação-rejeição é um método de geração de amostras aleatórias de uma distribuição desejada  $f$  a partir de uma distribuição facilitadora  $g$  que é mais fácil de ser gerada. O método funciona gerando amostras aleatórias de  $g$  e aceitando aquelas que satisfazem uma condição baseada na relação entre as duas distribuições. Aquelas amostras que não satisfazem a condição são rejeitadas e novas amostras são geradas até que uma amostra aceitável seja obtida.

Figura 18 – (a) Fase original(livre de ruído) (b) Fase ruidosa



Fonte: Fonte: O autor (2025).

Figura 19 – Antes e depois da filtragem: (a) Lee Refinado, (b) LInSARRFE, (c) TcNfilter, (d) TcCfilter. Fase desembrulhada após a aplicação do filtro: (e) Lee Refinado, (f) LInSARRFE, (g) TcNfilter, (h) TcCfilter



Fonte: O autor (2025).

A Tabela 3 apresenta as métricas RMSE, SSIM, MD e VD, comparando a fase pura com a imagem filtrada, além do tempo de execução para cada filtro. De acordo com os dados apresentados na Tabela 3, o TcCfilter se destaca em relação ao RMSE, apresentando o menor valor e a maior precisão entre os filtros analisados. Em seguida, os filtros LInSARRFE e

TcNfilter possuem o mesmo valor de RMSE, enquanto o filtro Lee Refinado apresenta um valor ligeiramente superior. Em relação ao SSIM, a métrica que avalia a preservação da estrutura da imagem, o filtro LInSARRFE é o mais eficaz, com um valor próximo de 1. Os filtros TcCfilter e Lee Refinado também apresentam valores de SSIM semelhantes, com uma leve vantagem para o TcCfilter.

Ao comparar os filtros propostos com o Lee Refinado, observamos que LInSARRFE e TcNfilter apresentam desempenho equivalente na preservação da fase original, com médias e variâncias das diferenças próximas às do Lee Refinado. Uma média baixa indica que os valores filtrados estão próximos aos originais, enquanto uma variância baixa reflete consistência nas diferenças e estabilidade nos dados. O TcNfilter se destaca por apresentar a menor variância entre os filtros analisado, garantindo maior estabilidade na preservação da fase. Por outro lado, o TcCfilter apresentou a maior média e variância das diferenças, indicando maior alteração na fase original e menor capacidade de preservação das informações interferométricas.

Em termos de processamento, os filtros propostos se destacam pela eficiência, apresentando os menores tempos de execução, tornando-os ideais para aplicações que exigem processamento rápido. Notavelmente, o TcNfilter alcança a maior economia de tempo em comparação com o filtro Lee Refinado, resultando em uma redução aproximada de 92% do tempo. Em contraste, o filtro Lee Refinado é o mais lento, o que pode limitar sua aplicação em contextos práticos. Portanto, os filtros propostos demonstram ser comparáveis ao filtro Lee Refinado em redução de ruídos e preservação de qualidade, e superiores em eficiência computacional.

Tabela 3 – Eficiência dos filtros para dados simulados

Filtro	RMSE	SSIM	MD(rad)	VD(rad)	Tempo de Execução(s)	Tempo Otimizado(%)
Lee Refinado	1,255	<b>0,687</b>	<b>0,212</b>	0,034	4790,76	—
LInSARRFE	1,269	0,644	0,213	0,034	1460,44	69,51
TcNfilter	1,297	0,670	0,213	<b>0,033</b>	<b>404,10</b>	91,56
TcCfilter	<b>1,251</b>	0,641	0,228	0,036	427,73	91,07

Os resultados apresentados nesta seção destacam a eficácia dos filtros propostos em comparação com o filtro Lee Refinado, tanto em precisão quanto em eficiência. O TcCfilter se destaca no quesito RMSE, sugerindo que ele preserva melhor a precisão dos dados, enquanto o LInSARRFE é o mais eficaz na preservação da estrutura da imagem, com o maior valor de SSIM. Além disso, se comparados com o Lee Refinado, os filtros propostos demonstram

uma vantagem significativa em termos computacionais, com o TcNfilter alcançando a maior economia de tempo. Essa combinação de alta precisão, preservação da qualidade da imagem e eficiência computacional posiciona os filtros LInSARRFE, TcNfilter e TcCfilter como opções adequadas para aplicações que exigem rapidez e qualidade no processamento de dados.

Com base nos dados da Tabela 3, o filtro proposto, TcNfilter, preserva as principais informações interferométricas, como a variação de fase (VD), enquanto mantém um excelente desempenho em termos de tempo de execução, proporcionando uma maior economia de tempo de processamento, sem comprometer significativamente a qualidade dos dados.

### 4.3 FILTRAGEM DE DADOS INSAR

Os sensores SAR de banda C, Sentinel-1A e Sentinel-1B da Agência Espacial Europeia (ESA - European Space Agency), fornecem aquisições de imagens SAR repetidas com uma taxa de observação previsível, proporcionando uma excelente base de dados para análises ambientais usando técnicas InSAR. Nesta tese, exploramos duas aplicações utilizando imagens SAR de banda C fornecidas pelo Sentinel-1.

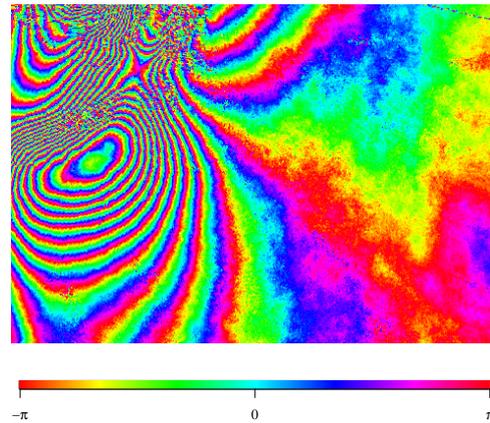
#### 4.3.1 La Cumbre

Na primeira aplicação, analisamos a erupção do vulcão La Cumbre, localizado na Ilha Fernandina, no arquipélago de Galápagos, no Equador, que ocorreu em janeiro de 2020. Recentemente, o vulcão La Cumbre entrou em erupção em 12 de janeiro de 2020, com 11 tremores sísmicos registrados naquele dia, atingindo uma magnitude máxima de 4,7 na escala Richter. A Região de Interesse (ROI) é ilustrada na Figura 20 e tem dimensões de  $346 \times 492$  pixels.

De modo análogo ao realizado na seção 4.2, aplicamos os filtros adotados nesta pesquisa ao interferograma ilustrado na Figura 20. As Figuras, 21(a), 21(b), 21(c) e 21(d) mostram os resultados após aplicação dos filtros, Lee InSAR Refinado, LInSARRFE, TcNfilter e TcCfilter a Figura 20. Enquanto, as Figuras 21(e), 21(f), 21(g) e 21(h) mostram as fases desembrulhadas correspondentes.

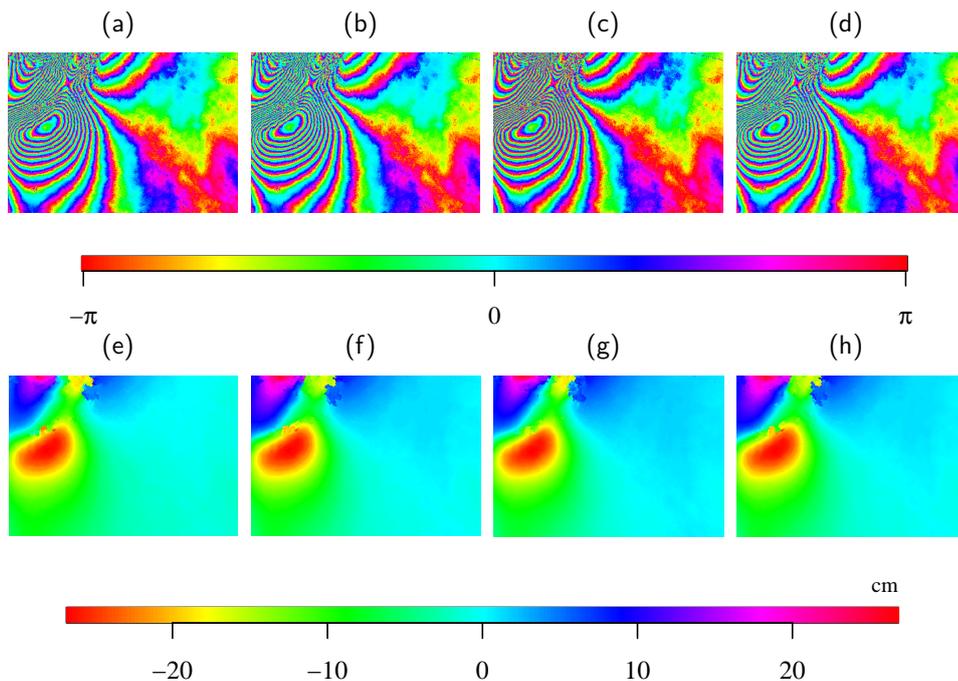
Com base nos dados da Tabela 4, podemos observar que o filtro Lee Refinado reduz significativamente o número de resíduos, demonstrando boa capacidade de filtragem. No entanto, apresenta o maior tempo de execução, o que pode limitar sua aplicabilidade em cenários que

Figura 20 – Imagem de diferença de fase ou interferograma



Fonte: (PECHNIKOV, 2020)

Figura 21 – Antes de depois da filtragem: (a) Lee InSAR Refinado, (b) LInSARRFE, (c) TcNfilter, (d) TcC-filter. Fase desembrulhada após aplicação do filtro: (e) Lee InSAR Refinado, (f) LInSARRFE, (g) TcNfilter, (h) TcCfilter



Fonte: O autor (2025).

exigem processamento rápido. O filtro LInSARRFE, por sua vez, apresenta a maior redução de resíduos entre os filtros avaliados, alcançando aproximadamente 82% em relação ao interferograma ilustrado na Figura 20, antes do processo de filtragem. Além disso, seu tempo de execução é significativamente menor, com uma redução de cerca de 79% no tempo de processamento.

Os filtros TcNfilter e TcCfilter, baseados em modelos empíricos, apresentam reduções de

resíduos um pouco superiores às do Lee Refinado. Por outro lado, destacam-se pela eficiência computacional. O TcNfilter é o mais eficiente, com uma redução em torno de 97% no tempo de execução, enquanto o TcCfilter também demonstra excelente desempenho, com uma economia de aproximadamente 96%, no tempo de processamento. Esses resultados indicam que os filtros propostos oferecem uma combinação eficaz de qualidade de filtragem e eficiência computacional, tornando-os alternativas viáveis ao filtro Lee Refinado em cenários que exigem rapidez e precisão.

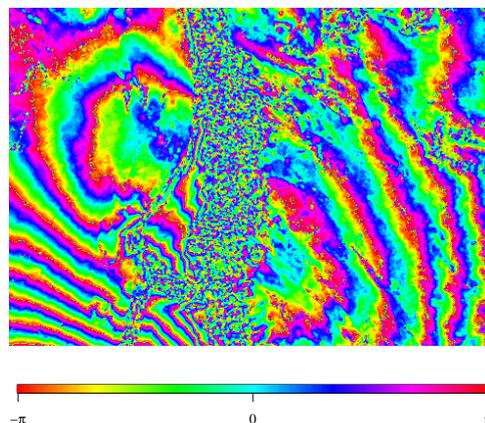
Tabela 4 – Eficiência dos filtros para dados SAR – La Cumbre

Filtro	Número de Resíduos	Redução de Resíduos(%)	Tempo de Execução(s)	Tempo Otimizado(%)
Não Filtrada	11 838	—	—	—
Lee Refinado	2481	79,04	57 851,37	—
LInSARRFE	<b>2145</b>	<b>81,88</b>	12 069,78	79,02
TcNfilter	2432	79,45	<b>2023,66</b>	96,50
TcCfilter	2428	79,49	2436,25	95,78

#### 4.3.2 Fronteira Irã-Iraque

Para esta aplicação, utilizamos uma imagem SAR, do Sentinel-1, com dimensões  $346 \times 492$  pixels. A região em estudo, conforme mostrado na Figura 22, abrange a área afetada por um terremoto ocorrido na fronteira Irã-Iraque, com a cidade curda iraquiana de Halabja, em 12 de novembro de 2017.

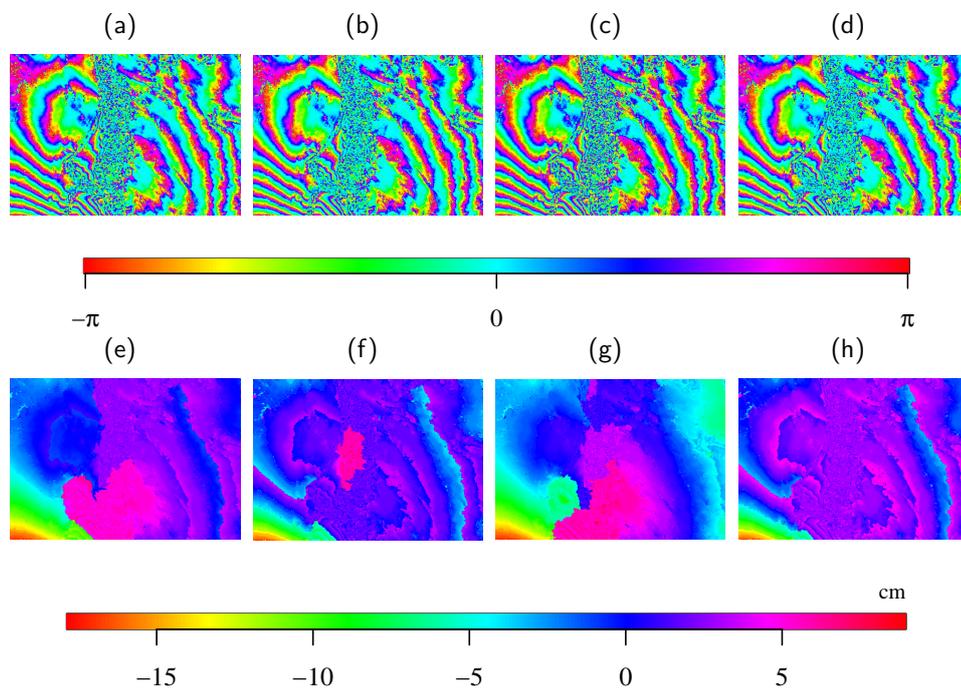
Figura 22 – Imagem de diferença de fase ou interferograma



Fonte: (PECHNIKOV, 2017)

Equivalente ao realizado anteriormente, nesta aplicação, além dos métodos propostos, aplicamos também o filtro Lee refinado ao interferograma mostrado na Figura 22. As Figuras 23(a),23(b),23(c) e 23(d) apresentam os resultados após a aplicação dos filtros Lee InSAR Refinado, LInSARRFE, TcNfilter e TcCfilter, respectivamente, ao interferograma da Figura 22. Ao passo que, as Figuras 23(e),23(f),23(g) e 23(h) exibem as fases desembrulhadas correspondentes."

Figura 23 – Comparação de resultados filtrados: (a) Lee InSAR Refinado, (b) LInSARRFE, (c) TcNfilter, (d) TcCfilter. Fase desembrulhada após aplicação do filtro: (e) Lee InSAR Refinado, (f) LInSARRFE, (g) TcNfilter, (h) TcCfilter



Fonte: O autor (2025).

Após o processo de filtragem, analisamos e comparamos a eficiência de cada filtro com base nas métricas de número de resíduos e tempo de processamento. Os resultados dessa análise estão apresentados na Tabela 5. Os dados exibidos na Tabela 5 mostram que os métodos propostos são amplamente superiores ao Lee Refinado com relação ao tempo de execução, com otimizações significativas que variam de 78 % a 94 %. Garantindo o equilíbrio entre a redução de resíduos e a eficiência computacional. Esses resultados reforçam o bom desempenho dos métodos propostos, oferecendo eficácia na qualidade da filtragem e eficiência computacional.

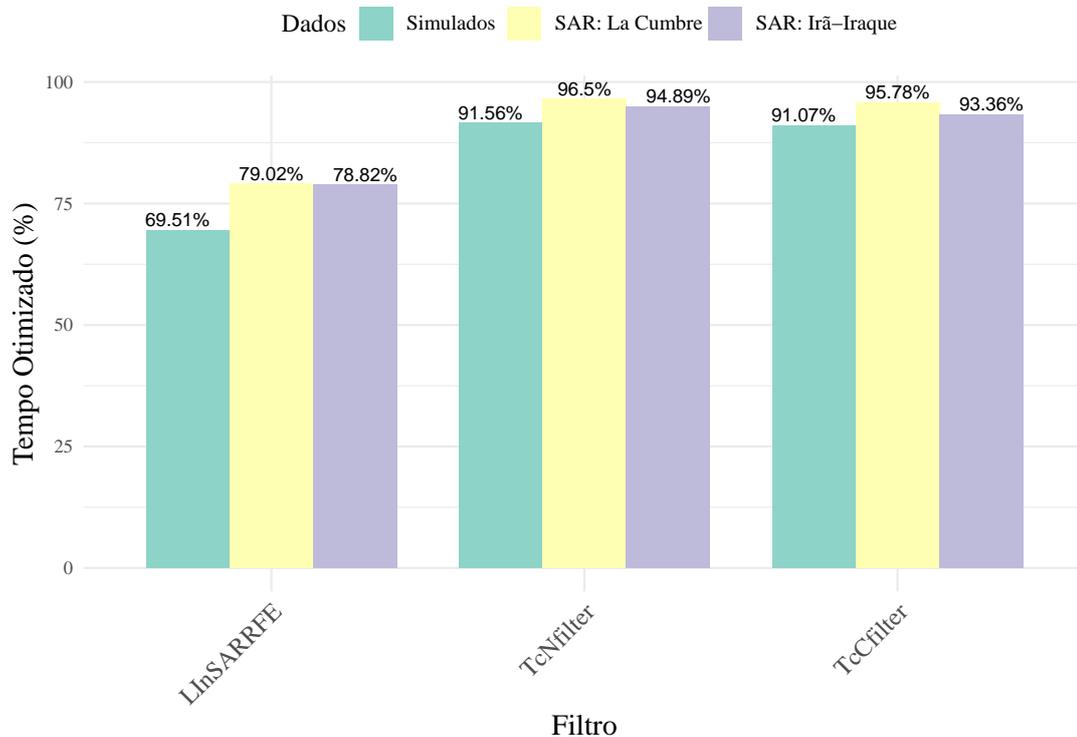
Tabela 5 – Eficiência dos filtros para dados SAR – Irã-Iraque

Filtro	Número de Resíduos	Redução de Resíduos(%)	Tempo de Execução(s)	Tempo Otimizado(%)
Não Filtrada	8054	—	—	—
Lee Refinado	4329	46,25	52 316,21	—
LInSARRFE	3832	38,32	11 078,52	78,82
TcNfilter	4839	39,91	<b>2671,41</b>	94,89
TcCfilter	<b>3682</b>	54,38	3470,81	93,36

#### 4.4 DESEMPENHO TEMPORAL DOS FILTROS PROPOSTOS EM COMPARAÇÃO AO LEE REFINADO

O gráfico representado na Figura 24 fornece informações sobre a eficiência computacional do LInSARRFE, TcNfilter e TcCfilter, filtros propostos nesta pesquisa, em comparação com o Lee Refinado. Os resultados apresentados no gráfico em 24, refletem a avaliação dos filtros propostos em dois cenários: em dados simulados e em dados SAR.

Figura 24 – Tempo Otimizado



Fonte: O autor (2025)

Os impactos dos filtros propostos são altamente positivos, apresentando-se como alternativas eficazes e eficientes com relação ao tempo de processamento, economizando entre aproximadamente 70 % e 97 % do tempo em comparação com o Lee Refinado. Isso implica que o LInSARRFE, o TcNfilter e o TcCfilter são escolhas eficientes e práticas para aplicações que exigem filtragem de ruído em dados InSAR, especialmente quando se considera a necessidade de processar grandes volumes de dados em um tempo limitado.

## 5 CONSIDERAÇÕES FINAIS

Esta tese desenvolveu uma análise abrangente de modelos probabilísticos, físicos e empíricos, aplicados à análise de dados de fase, destacando suas principais características, as relações entre eles e enfatizando novos modelos e funcionalidades. Além disso, desenvolveu, avaliou e disponibilizou os pacotes `FDDPhase` e `Filters`. O primeiro destinado a análise estatística para dados de fase, enquanto o segundo abrange métodos de filtragem para o ruído de fase.

Durante o desenvolvimento desta tese, uma etapa fundamental envolveu a proposição e implementação dos filtros `LInSARRFE`, `TcNfilter` e `TcCfilter` (disponíveis no pacote `Filters`). Estes filtros foram concebidos com base nos modelos em (3.21), definido por Gierull (2021),  $\mathcal{NT}$  e  $\mathcal{CT}$ , respectivamente. O objetivo principal desses métodos de filtragem é mitigar o ruído de fase, simultaneamente assegurando a preservação da qualidade da imagem. Para avaliar o desempenho desses filtros, foram utilizadas métricas essenciais, tais como RMSE, SSIM, MD, VD, número de resíduos e Tempo de Execução. Estas métricas desempenharam um papel fundamental ao quantificar a eficácia dos filtros propostos.

Para avaliar a eficiência dos filtros propostos utilizamos três aplicações, uma com dados simulados e outra duas com dados SAR fornecidas pelo Sentinel-1, banda C. Primeiramente, utilizamos uma imagem simulada com dimensões de  $128 \times 128$  pixels. Em seguida, analisamos a erupção do vulcão La Cumbre, localizado na Ilha Fernandina, no arquipélago de Galápagos, no Equador, que ocorreu em 2020. Na terceira aplicação, a região estudada abrange a área afetada pelo terremoto ocorrido em 2017, na fronteira entre Irã e Iraque, próximo à cidade curda iraquiana de Halabja.

Na continuação da pesquisa, realizamos uma análise comparativa da eficiência, tanto no processo de filtragem quanto no aspecto computacional, entre os filtros propostos e o Lee Refinado. De maneira geral, os filtros propostos se destacam como uma escolha eficaz tanto em precisão quanto em tempo de processamento, apresentando equilíbrio entre qualidade da filtragem e eficiência computacional. Esse equilíbrio resulta em economias de tempo que variam de aproximadamente 70% a 97%, evidenciando que, além de manterem a qualidade da filtragem, os filtros propostos oferecem uma vantagem significativa em termos de desempenho computacional.

Portanto, com base nas análises realizadas ao longo da pesquisa e apresentadas nesta tese, este trabalho representa uma contribuição relevante para o avanço do processamento de

imagens SAR. Apresentamos modelos teóricos para a diferença de fase e propusemos filtros eficazes e eficientes para a redução de ruído em dados interferométricos. As análises teóricas e os experimentos em dados simulados e SAR validam a robustez e relevância da proposta.

Ao considerar as descobertas desta pesquisa, emergem diversas perspectivas para futuras investigações. Uma área de foco sugerida é aprofundar a análise da convergência dos modelos físicos para diferença de fase. Por exemplo, para valores reais não inteiros de  $L$ , o modelo definido em (3.20) funciona, por outro lado, a expressão dada em (3.21), obviamente não. No entanto, é provável que se possa mostrar que uma interpolação linear ou quadrática dos valores de fase entre dois ou mais valores inteiros adjacentes também pode funcionar muito bem.

Além disso, explorar e desenvolver novos modelos empíricos, adaptáveis a diferentes cenários InSAR, representa uma abordagem promissora. A integração de técnicas de aprendizado de máquina surge como uma direção promissora, visando desenvolver modelos capazes de aprender padrões específicos nas imagens SAR e ajustar automaticamente os parâmetros dos filtros, representando um avanço significativo na eficácia dessas abordagens.

## REFERÊNCIAS

- ALCOFORADO, L. F. *Utilizando a linguagem R: Conceitos, manipulação, visualização modelagem e elaboração de relatórios*. [S.l.]: Alta Books, 2021. v. 1. ISBN 9788550814421.
- ARFKEN, G.; WEBER, H. *Mathematical methods for physicists*. [S.l.]: CAMPUS - GRUPO ELSEVIER; Ciencias Exatas edition, 2007. v. 6. 936 p. ISBN 9788535220506.
- ARGENTI, F.; LAPINI, A.; BIANCHI, T.; ALPARONE, L. A tutorial on speckle reduction in synthetic aperture radar images. *IEEE Geoscience and Remote Sensing Magazine*, v. 1, n. 3, p. 6–35, 2013.
- ARTHUR, P.; MARKUS, N.; GRAEME, D. R. *Circular Statistics in R*. New York, United States: Oxford, 2013. 208 p. ISBN 9780199671137.
- ATEYA, S.; MADHAGI, E. On multivariate truncated generalized Cauchy distribution. *Statistical Papers*, v. 54, 08 2013.
- BAIER, G.; ROSSI, C.; M., L.; ZHU, X. X.; MEMLER, R. A nonlocal InSAR filter for high-resolution DEM generation from TanDEM-X interferograms. *IEEE Trans. Geosci. Remote Sens.*, v. 56, n. 11, p. 6469–6483, nov 2018.
- BORBA, A. A. *Fusão de evidências de bordas dos canais de intensidade de imagem de radar polarimétrico de abertura sintética*. Tese (Tese) — Universidade Presbiteriana Mackenzie, 2020. Disponível em: <<http://tede.mackenzie.br/jspui/handle/tede/4716>>.
- CHAO, C.-F.; CHEN, K.-S.; LEE, J.-S. Refined filtering of interferometric phase from InSAR data. *IEEE Trans. Geosci. Remote Sens.*, v. 51, n. 12, p. 5315–5323, 2013.
- CREMERS, J.; KLUGKIST, I. One direction? A tutorial for circular data analysis using R with examples in cognitive psychology. *Front Psychol*, 2018.
- DENG, X.; LÓPEZ-MARTÍNEZ, C.; CHEN, J.; HAN, P. Statistical modeling of polarimetric SAR data: A survey and challenges. *Remote Sensing*, v. 9, n. 4, 2017. ISSN 2072–4292. Disponível em: <<https://www.mdpi.com/2072-4292/9/4/348>>.
- DOMG, Y.; MILNE, A.; FORSTER, B. Toward edge sharpening: A SAR speckle filtering algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, v. 39, n. 4, p. 851–863, 2001.
- DOWNS, T. D. Some relationships among the von Mises distributions of different dimensions. *Biometrika*, Oxford University Press, Biometrika Trust, v. 53, n. 1/2, p. 269–272, 1966. ISSN 00063444. Disponível em: <<http://www.jstor.org/stable/2334080>>.
- DOWNS, T. D.; GOULD, A. L. Some relationships between the normal and von Mises distributions. *Biometrika*, v. 54, n. 3-4, p. 684–687, 12 1967.
- FRERY, A.; MULLER, H.-J.; YANASSE, C.; SANT'ANNA, S. A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*, v. 35, n. 3, p. 648–659, 1997.
- FRERY, A. C.; NASCIMENTO, A. D. C.; CINTRA, R. J. Analytic expressions for stochastic distances between relaxed complex Wishart distributions. *IEEE Trans. Geosci. Remote Sens.*, v. 52, n. 2, p. 1213–1226, 2014.

- GAMA, F.; PARADELLA, W.; MURA, J.; SANTOS, A. Surface deformation radar interferometry techniques using orbital data. *Brazilian Symposium on Remote Sensing*, Foz do Iguaçu, abr 2013.
- GAO, G. Statistical modeling of SAR images: A survey. *Sensors*, v. 10, p. 775–795, 2010.
- GIERULL, C. H. Closed-form expressions for InSAR sample statistics and its application to non-gaussian data. *IEEE Trans. Geosci. Remote Sens.*, v. 59, n. 5, p. 3967–3980, 2021.
- GOODMAN, N. R. Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction). *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 34, n. 1, p. 152–177, 1963. Disponível em: <<https://doi.org/10.1214/aoms/1177704250>>.
- HAYNES, M. S. Fast and accurate computation of the multilook interferometric phase probability density function. *IEEE Geosci. Remote Sens. Lett.*, v. 14, n. 5, p. 769–772, 2017.
- HORTA, M. M. *Modelos de mistura de distribuições na segmentação de imagens SAR polarimétricas multilook*. Tese (Tese) — Instituto de Física de São Carlos, 2009. Disponível em: <<https://teses.usp.br/teses/disponiveis/76/76132/tde-06072009-215138/pt-br.php>>.
- I., F. N. *Statistical Analysis of Circular Data*. Cambridge: Cambridge University Press, 1995.
- JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. *Continuous Univariate Distributions*. 2nd. ed. New York: John Wiley & Sons, 1994. (Wiley Series in Probability and Statistics). A Wiley-Interscience Publication.
- JONES, C. E.; P, M.; RAO, S. The NISAR mission's capabilities for natural hazards monitoring. In: *IEEE Int. Geosci. and Remote Sens. Symp. IGARSS*. [S.l.: s.n.], 2021. p. 1711–1714.
- JPL, N. *Interferometry*. n.d. <<https://nisar.jpl.nasa.gov/mission/get-to-know-sar/interferometry/#:~:text=In%20an%20interferogram%2C%20the%20two,the%20spread%20of%20the%20fringes>>. Accessed on: 21 October 2024.
- KATO, S.; JONES, M. C. A family of distributions on the circle with links to, and applications arising from, möbius transformation. *Journal of the American Statistical Association*, v. 105, n. 489, p. 249–262, March 2010. Disponível em: <<https://oro.open.ac.uk/25815/>>.
- KUAN, D.; SAWCHUK, A.; STRAND, T.; CHAVEL, P. Adaptive restoration of images with speckle. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 35, n. 3, p. 373–383, 1987.
- KURZ, G.; GILITSCHENSKI, I.; HANEBECK, U. Efficient evaluation of the probability density function of a wrapped normal distribution. In: *Proceedings of the 17th International Conference on Information Fusion*. [S.l.: s.n.], 2014. p. 1–7.
- LEE, J. S. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2, p. 165–168, 1980.
- LEE, J. S.; HOPPEL, K. W.; MANGO, S. A.; MILLER, A. R. Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery. *IEEE Trans. Geosci. Remote Sens.*, v. 32, n. 5, p. 1017–1028, set. 1994.

- LEE, J.-S.; PAPATHANASSIOU, K.; AINSWORTH, T.; GRUNES, M.; REIGBER, A. A new technique for noise filtering of SAR interferometric phase images. *IEEE Transactions on Geoscience and Remote Sensing*, v. 36, n. 5, p. 1456–1465, 1998.
- LEE, J. S.; POTTIER, E. *Polarimetric Radar Radar Imaging*. New York: CRC Press, 2009. ISBN 9781420054972.
- LEE, J.-S.; WEN, J.-H.; AINSWORTH, T.; CHEN, K.-S.; CHEN, A. Improved sigma filter for speckle filtering of SAR imagery. *IEEE Trans. Geosci. Remote Sens.*, v. 47, n. 1, p. 202–213, 2009.
- LI, T.; CHEN, K.-S.; LEE, J.-S. Enhanced interferometric phase noise filtering of the refined InSAR filter. *IEEE Geosci. Remote Sens. Lett.*, v. 17, n. 9, p. 1528–1532, 2020.
- MA, X.; WU, P.; WU, Y.; SHEN, H. A review on recent developments in fully polarimetric SAR image despeckling. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 11, n. 3, p. 743–758, 2018.
- MARDIA, K.; TAYLOR, C.; SUBRAMANIAM, G. Protein bioinformatics and mixtures of bivariate von Mises distributions for angular data. *Biometrics*, v. 63, n. 2, p. 505–512, June 2006.
- MARDIA, K. V. Statistics of directional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, [Royal Statistical Society, Wiley], v. 37, n. 3, p. 349–393, 1975.
- MARDIA K. V. AND JUPP, P. *Directional Statistics*. New York, United States: Wiley, 2000. 460 p. ISBN 9780471953333.
- MARTÍNEZ, C. L.; POTTIER, E. *Basic Principles of SAR Polarimetry*. In: Hajnsek, I., Desnos, YL. (eds) *Polarimetric Synthetic Aperture Radar. Remote Sensing and Digital Image Processing*. Gewerbestrasse 11, 6330 Cham, Switzerland: Springer, 2021. v. 25.
- MIGUEL, A. H.; DAVID, R. B.; MICHAEL, J. L.; MUNTHER, A. G. Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path. *Appl. Opt.*, Optica Publishing Group, v. 41, n. 35, p. 7437–7444, Dec 2002. Disponível em: <<https://opg.optica.org/ao/abstract.cfm?URI=ao-41-35-7437>>.
- MURA, J. C. *Automatic geocoding of interferometric synthetic aperture radar images: Geo-InSAR system*. Tese (Tese) — Instituto Nacional de Pesquisas Espaciais (INPE), 2000.
- MURA, J. C. Coherence optimization of interferometric and polarimetric SAR images in the P-band in forest areas. *Brazilian Symposium on Remote Sensing*, Goiânia, abr 2005.
- NADARAJAH, S.; KOTZ, S. R programs for computing truncated distributions. *Journal of Statistical Software*, American Statistical Association, v. 16, n. Code Snippet 2, p. 1–8, August 2006. Published by the American Statistical Association. Disponível em: <<http://www.jstatsoft.org/v16/c02>>.
- NADARAJAH, S.; KOTZ, S. A truncated Cauchy distribution. *International Journal of Mathematical Education in Science and Technology*, v. 37, p. 605–608, 07 2006.
- NASCIMENTO, A. D. C.; FRERY, A. C.; CINTRA, R. J. Detecting changes in fully polarimetric SAR imagery with statistical information theory. *IEEE Transactions on Geoscience and Remote Sensing*, v. 57, n. 3, p. 1380–1392, 2019.

- NOVO, M. L. M. *Sensoriamento remoto: Princípios e aplicações*. São Paulo: Blucher, 2010. ISBN 9788521205401.
- PARADELLA, W. R.; MURA, J. C.; GAMA, F. G.; SANTOS, A. R.; SILVA, G. G. Radar imagers SAR orbitals: trends in systems and applications. *Simpósio Brasileiro de Sensoriamento Remoto*, João Pessoa, abr 2015.
- PARADELLA, W. R.; SANTOS, A. R.; VENEZIANI, P.; CUNHA, E. S. P. Imaging radars in geosciences: Status and perspectives. *Proceedings XII Brazilian Symposium on Remote Sensing*, Goiânia, abr 2005.
- PAULA, F. V.; NASCIMENTO, A. D. C.; AMARAL, G. J. A.; CORDEIRO, G. M. Generalized cardioid distributions for circular data analysis. *Stats*, v. 4, n. 3, p. 634–649, 2021. ISSN 2571-905X. Disponível em: <<https://www.mdpi.com/2571-905X/4/3/38>>.
- PECHNIKOV, A. *PyGMTSAR Co-Seismic Interferogram: Iran–Iraq Earthquake, 2017*. 2017. <<https://github.com/AlexeyPechnikov/pygmtsar/tree/pygmtsar2/notebooks>>. Accessed on: 18 March 2024.
- PECHNIKOV, A. *PyGMTSAR Co-Seismic Interferogram: La Cumbre Volcano Eruption, 2020*. 2020. <<https://github.com/AlexeyPechnikov/pygmtsar/tree/pygmtsar2/notebooks>>. Accessed on: 10 March 2024.
- PU, L.; ZHANG, X.; ZHOU, Z.; SHI, J.; WEI, S.; ZHOU, Y. A phase filtering method with scale recurrent networks for InSAR. *Remote Sensing*, v. 12, n. 20, 2020. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/12/20/3453>>.
- RAO, J. S.; SENGUPTA, A. *Topics in Circular Statistics*. River Edge, N.J: World Scientific, 2001. v. 5. 336 p. (Series on Multivariate Analysis, v. 5). ISBN 9710237782.
- ROBERT, C. P. Simulation of truncated normal variables. *Statistics and Computing*, v. 5, n. 2, p. 121–125, 1995.
- SICA, F.; GOBBI, G.; RIZZOLI, P.; BRUZZONE, L.  $\phi$ -Net: Deep residual learning for InSAR parameters estimation. *IEEE Trans. Geosci. Remote Sens.*, v. 59, n. 5, p. 3917–3941, 2021.
- SINGH, A. *NASA’s NISAR Project: Interview with Scientist Paul Rosen*. The Week, 2023. Accessed: 2024-06-21. Disponível em: <<https://www.theweek.in/theweek/specials/2023/06/24/nasa-s-nisar-project-scientist-paul-rosen-interview.html>>.
- SINTES, C.; COURMONTAGNE, P.; LLORT-PUJOL, G.; CAILLEC, J.-M. L. Gaussian approximation of interferometric PDF for MLE derivation. In: *2012 Oceans*. [S.l.: s.n.], 2012. p. 1–5.
- SUN, X.; ZIMMER, A.; MUKHERJEE, S.; KOTTAYIL, N. K.; GHUMAN, P.; CHENG, I. DeepinSAR— A deep learning framework for SAR interferometric phase restoration and coherence estimation. *Remote Sensing*, v. 12, n. 14, 2020. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/12/14/2340>>.
- VELHO, L. F.; NETO, J. A.; SIMÕES, J. C. Utilização do filtro lee na redução do speckle em imagens SAR usadas na determinação da velocidade de geleiras. *Revista Brasileira de Cartografia*, v. 63, n. 2, 2012. Disponível em: <<https://seer.ufu.br/index.php/revistabrasileiracartografia/article/view/43737>>.

---

VITALE, S.; FERRAIOLI, G.; PASCAZIO, V.; SCHIRINZI, G. InSAR-MONet: Interferometric SAR phase denoising using a multiobjective neural network. *IEEE Trans. Geosci. Remote Sens.*, Institute of Electrical and Electronics Engineers (IEEE), v. 60, p. 1–14, 2022. ISSN 1558-0644.

WANG, C.; CHANG, L.; WANG, X.-S.; ZHANG, B.; STEIN, A. Interferometric synthetic aperture radar statistical inference in deformation measurement and geophysical inversion: A review. *IEEE Geosci. Remote Sens. Mag.*, v. 12, n. 1, p. 8–35, 2024.

ZEYADA, H. H.; MOSTAFA, M. S.; EZZ, M. M.; NASR, A. H.; HARB, H. M. Resolving phase unwrapping in interferometric synthetic aperture radar using deep recurrent residual u-net. *The Egyptian Journal of Remote Sensing and Space Science*, v. 25, n. 1, p. 1–10, 2022. ISSN 1110-9823.

ZONATTA, D. C.; FERREIRA, M. P.; ZORTEA, M. *Processamento de imagens de satélite*. [S.l.]: Oficinas de textos: São Paulo, 2019. v. 1. ISBN 9788579753169.

## APÊNDICE A – FUNÇÕES DE BESSEL

As funções de Bessel, primeiro definidas pelo matemático Daniel Bernoulli<sup>1</sup> e depois generalizadas por Friedrich Bessel<sup>2</sup>, são soluções canônicas da equação de Bessel de ordem  $\nu$  dada pela Equação Diferencial Ordinária (EDO) a seguir:

$$x^2 y'' + xy' + (x^2 - \nu^2)y = 0,$$

onde  $\nu$  é um real não negativo.

O ponto  $x_0 = 0$  é um ponto singular regular para a equação de Bessel, então, aplicando o método de Frobenius<sup>3</sup> existe pelo menos uma solução da forma:

$$y = x^r \sum_{n=0}^{\infty} C_n x^n = \sum_{n=0}^{\infty} C_n x^{n+r}, \quad C_0 \neq 0.$$

Derivando  $y$  duas vezes em relação a  $x$ , obtemos:

$$\frac{dy}{dx} = \sum_{n=0}^{\infty} (n+r) C_n x^{n+r-1}, \quad \frac{dy^2}{dx} = \sum_{n=0}^{\infty} (n+r)(n+r-1) C_n x^{n+r-2}.$$

Agora, substituindo o valor de  $y$  apresentado acima na equação de Bessel, obtemos:

$$\sum_{n=0}^{\infty} (n+r)(n+r-1) C_n x^{n+r} + \sum_{n=0}^{\infty} (n+r) C_n x^{n+r} + \sum_{n=0}^{\infty} C_n x^{n+r+2} - \sum_{n=0}^{\infty} \nu^2 C_n x^{n+r} = 0.$$

Agrupando os termos dos somatórios 1, 2 e 4 da última igualdade, temos:

$$\sum_{n=0}^{\infty} [(n+r)^2 - \nu^2] C_n x^{n+r} + \sum_{n=0}^{\infty} C_n x^{n+r+2} = 0.$$

No segundo somatório da igualdade anterior, fazendo  $k = n + 2$  e fatorando o coeficiente do primeiro somatório, obtemos:

$$\sum_{n=0}^{\infty} (n+r+\nu)(n+r-\nu) C_n x^{n+r} + \sum_{k=2}^{\infty} C_{k-2} x^{k+r} = 0.$$

No segundo somatório da última igualdade, substituindo o índice  $k$  por  $n$ , separando os dois primeiros termos do primeiro somatório, obtemos:

- 
- <sup>1</sup> Daniel Bernoulli foi um matemático e físico suíço e foi um dos muitos matemáticos proeminentes da família Bernoulli de Basel. Ele é particularmente lembrado por suas aplicações da matemática à mecânica, especialmente a mecânica dos fluidos, e por seu trabalho pioneiro em probabilidade e estatística. Acesso em: 15/07/2021. Disponível em: [https://en.wikipedia.org/wiki/Daniel\\_Bernoulli](https://en.wikipedia.org/wiki/Daniel_Bernoulli)
  - <sup>2</sup> Friedrich Wilhelm Bessel foi um alemão astrônomo, matemático, físico e geodesta. Ele foi o primeiro astrônomo a determinar valores confiáveis para a distância do Sol a outra estrela. Acesso em: 15/07/2021. Disponível em: [https://en.wikipedia.org/wiki/Friedrich\\_Bessel](https://en.wikipedia.org/wiki/Friedrich_Bessel)
  - <sup>3</sup> Ferdinand Georg Frobenius foi um matemático alemão, mais conhecido por suas contribuições à teoria das funções elípticas, equações diferenciais, teoria dos números e teoria dos grupos. Acesso em: 21/07/2021. Disponível em: [https://pt.wikipedia.org/wiki/Ferdinand\\_Georg\\_Frobenius](https://pt.wikipedia.org/wiki/Ferdinand_Georg_Frobenius)

$$(r + \nu)(r - \nu)C_0x^r + (r + \nu + 1)(r - \nu + 1)C_1x^{r+1} + \sum_{k=2}^{\infty} [(n + r + \nu)(n + r - \nu)C_n + C_{n-2}]x^{n+r} = 0.$$

Por hipótese  $C_0 \neq 0$ , então:

$$(r + \nu)(r - \nu) = 0 \quad (\text{A.1})$$

$$(r + \nu + 1)(r - \nu + 1)C_1 = 0 \quad (\text{A.2})$$

$$(n + r + \nu)(n + r - \nu)C_n + C_{n-2} = 0 \quad (\text{A.3})$$

A igualdade em (A.1) é a equação indicial. Suas raízes são  $r_1 = \nu$  e  $r_2 = -\nu$ . Embora  $\nu$  e  $-\nu$  produzam a mesma EDO, é convencional definir diferentes funções de Bessel para esses dois vetores de forma que as funções de Bessel sejam principalmente funções suaves de  $\nu$ .

**Primeira solução:** Para  $r_1 = \nu \geq 0$ , equação em (A.2) se torna,

$$(2\nu + 1)C_1 = 0,$$

ou seja,

$$C_1 = 0. \quad (\text{A.4})$$

A relação de recorrência em (A.3) torna-se,

$$(n + 2\nu)nC_n + C_{n-2} = 0, \quad n = 2, 3, 4, \dots \quad (\text{A.5})$$

De (A.5) segue que,

$$C_n = -\frac{C_{n-2}}{n(n + 2\nu)}, \quad n = 2, 3, 4, \dots \quad (\text{A.6})$$

Agora, dos resultados em (A.4) e (A.6), e deixando para escolher mais tarde o valor de  $C_0$ , obtemos  $C_{2n+1} = 0$  e

$$\begin{aligned} C_2 &= -\frac{C_0}{2(2 + 2\nu)} \\ C_4 &= \frac{C_0}{2 \cdot 4(2 + 2\nu)(4 + 2\nu)} \\ C_6 &= -\frac{C_0}{2 \cdot 4 \cdot 6(2 + 2\nu)(4 + 2\nu)(6 + 2\nu)} \\ &\vdots \end{aligned}$$

ou seja,

$$\begin{aligned} C_2 &= \frac{(-1)^n C_0}{1(1+\nu) \cdot 2^2} \\ C_4 &= \frac{C_0}{1 \cdot 2(1+\nu)(2+\nu) \cdot 2^4} \\ C_6 &= -\frac{C_0}{1 \cdot 2 \cdot 3(1+\nu)(2+\nu)(3+\nu) \cdot 2^6} \\ &\vdots \end{aligned}$$

Portanto, de modo geral, temos

$$C_{2n} = -\frac{C_0}{n!(1+\nu)(2+\nu) \cdots (n+\nu) \cdot 2^{2n}}. \quad (\text{A.7})$$

Para obter as funções de Bessel de primeira espécie de ordem  $\nu$ , fazemos,

$$C_0 = \frac{1}{2^\nu \Gamma(1+\nu)}, \quad (\text{A.8})$$

onde,  $\Gamma(1+\nu)$  é a função Gama de  $(1+\nu)$ . Por definição,

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt, \quad x > 0.$$

Utilizando repetidas vezes a propriedade da função Gama dada por:

$$\Gamma(x+1) = x\Gamma(x),$$

obtemos,

$$\begin{aligned} (n+\nu) \cdots (2+\nu)(1+\nu)\Gamma(1+\nu) &= (n+\nu) \cdots (2+\nu)\Gamma(2+\nu) \\ &= (n+\nu) \cdots (3+\nu)\Gamma(3+\nu) \\ &= \cdots = \Gamma(n+\nu+1). \end{aligned}$$

Portanto, escolhendo  $C_0$  como em (A.8), (A.7) se torna,

$$C_{2n} = -\frac{(-1)^n}{n!\Gamma(n+\nu+1)2^{2n+\nu}}.$$

Obtemos, finalmente a solução:

$$y_1 = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\Gamma(n+\nu+1)} \left(\frac{x}{2}\right)^{2n+\nu}.$$

Esta função é chamada Função de Bessel de primeira espécie de índice  $\nu$  e denotada por  $J_\nu(x)$ , ou seja,

$$J_\nu(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\Gamma(n+\nu+1)} \left(\frac{x}{2}\right)^{2n+\nu}. \quad (\text{A.9})$$

Para  $\nu = m = 0, 1, 2, 3, \dots$ , podemos escrever,

$$J_m(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+m)!} \left(\frac{x}{2}\right)^{2n+m}, \quad (\text{A.10})$$

pois, por definição,  $\Gamma(n+m+1) = [(n+m+1)-1]! = (n+m)!$ .

De acordo com Rao e SenGupta (2001), para valores inteiros e positivo de  $r$ , podemos representar (A.9) usando uma integral:

$$J_r(x) = \frac{1}{\pi} \int_0^\pi \cos(r\theta - x \sin \theta) d\theta. \quad (\text{A.11})$$

Par Arfken e Weber (2007), a igualdade em (A.11) ocorre devido as propriedades de ortogonalidade das funções seno e cosseno.

Para  $\nu < 0$ , a equação em (A.10) resulta em,

$$J_{-\nu}(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n-\nu)!} \left(\frac{x}{2}\right)^{2n-\nu}.$$

Uma vez que  $\nu$  é um inteiro,  $(n-\nu)! \rightarrow \infty$ , para  $\nu = 0, \dots, (\nu-1)$ . Desta forma, podemos considerar que a série começa com  $n = \nu$ . Substituindo  $n$  por  $n + \nu$ , obtemos:

$$J_{-\nu}(x) = \sum_{n=0}^{\infty} \frac{(-1)^{n+\nu}}{n!(n+\nu)!} \left(\frac{x}{2}\right)^{\nu+2n}.$$

que mostra de forma imediata que  $J_\nu(x)$  e  $J_{-\nu}(x)$  não são independentes, mas estão relacionadas por:

$$J_{-\nu}(x) = (-1)^\nu J_\nu(x),$$

e

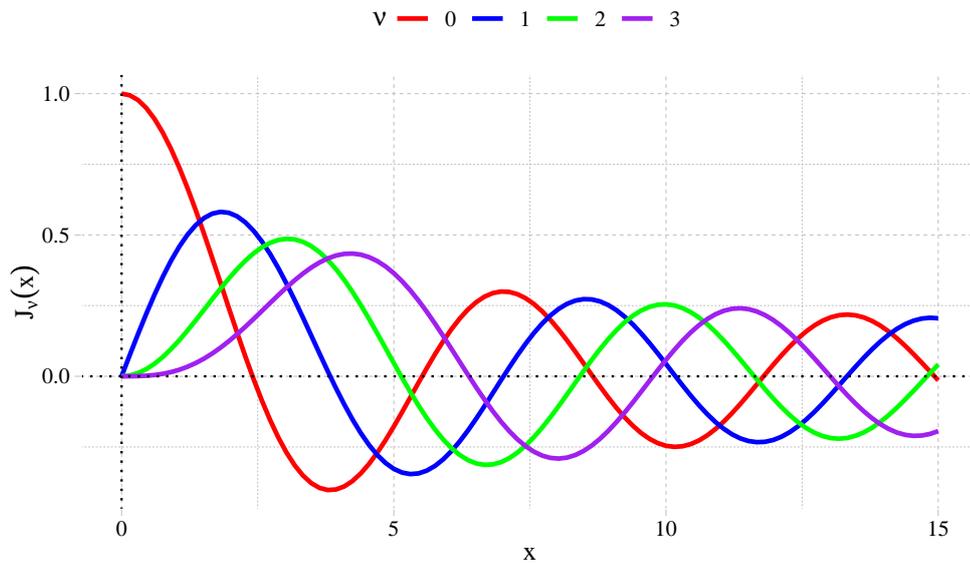
$$\frac{d}{dx} \{x^\nu J_\nu(x)\} = x^\nu J_{\nu+1}(x), \quad \nu > 0.$$

Em particular,

$$\begin{aligned} J_0(x) &= \frac{1}{2\pi} \int_0^{2\pi} e^{ix \cos \theta} d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \cos(x \cos \theta) d\theta \\ &= \frac{2}{\pi} \int_0^{\pi/2} \cos(x \cos \theta) d\theta \end{aligned}$$

e

$$\frac{d}{dx} (xJ_1(x)) = xJ_0(x).$$

Figura 25 – Função de Bessel de primeira espécie:  $J_\nu(x)$ 

Fonte: O autor (2025).

Os gráficos das funções de Bessel (Figura 25) parecem mais ou menos como funções oscilantes de seno ou cosseno que decaem proporcionalmente a  $x^{-\frac{1}{2}}$ , embora suas raízes não sejam geralmente periódicas, exceto assintoticamente para  $x$  grande.

A equação de Bessel não muda se substituirmos  $\nu$  por  $-\nu$ . Como visto anteriormente  $J_\nu(x)$  e  $J_{-\nu}(x)$  são soluções da equação de Bessel de índice  $\nu$ , neste sentido se faz necessário investigar em que casos estas duas funções são linearmente independentes. A seguir apresentamos algumas propriedades que nos ajudaram nesta análise.

1. Se  $\nu$  não é um inteiro, então  $J_\nu(x)$  e  $J_{-\nu}(x)$  são duas soluções linearmente independentes da equação de Bessel de índice  $\nu$ . Ou seja, não são múltiplas uma da outra e, portanto, são linearmente independentes.
2. Como visto anteriormente, para  $m = 0, 1, 2, 3, \dots$ ,

$$J_{-m}(x) = (-1)^m J_m(x).$$

Em resumo,

- Se  $\nu$  não é um inteiro, então  $J_\nu(x)$  e  $J_{-\nu}(x)$  são duas soluções linearmente independente da equação de Bessel de índice  $\nu$ ;
- Se  $\nu = m = 0, 1, 2, 3, \dots$ , então  $J_\nu(x)$  e  $J_{-\nu}(x)$  são linearmente dependentes, mais precisamente  $J_{-m}(x) = (-1)^m J_m(x)$ .

Para  $\nu = m = 0, 1, 2, 3, \dots$ , precisamos encontrar uma segunda solução linearmente independente de  $J_m(x)$ . Qualquer solução da equação de Bessel de índice  $m$ , linearmente independente de  $J_m(x)$ , é dita uma Função de Bessel de Segunda Espécie de Índice  $\nu$ .

Segundo Rao e SenGupta (2001), a função de Bessel modificada de primeiro tipo e ordem  $\nu$ , às vezes também chamada função de Bessel de argumento puramente imaginário, é dada por:

$$I_\nu(x) = i^{-\nu} J_\nu(ix). \quad (\text{A.12})$$

A Normalização de  $i^{-\nu}$  cancela os  $i^\nu$  de cada termo e transforma  $I_\nu(x)$  em real. Analisando a identidade em (A.12), obtemos:

$$\begin{aligned} I_\nu(x) &= i^{-\nu} J_\nu(ix) \\ &= \frac{1}{2\pi} \int_0^{2\pi} \cos(\nu\theta) e^{x \cos \theta} d\theta \\ &= \sum_{n=0}^{\infty} \frac{1}{n! \Gamma(n + \nu + 1)} \left(\frac{x}{2}\right)^{2n+\nu} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!(n + \nu)!} \left(\frac{x}{2}\right)^{2n+\nu}. \end{aligned}$$

De modo análogo, para  $\nu < 0$ , temos:

$$I_{-\nu}(x) = \sum_{n=0}^{\infty} \frac{1}{n!(n - \nu)!} \left(\frac{x}{2}\right)^{2n-\nu}.$$

Em particular,

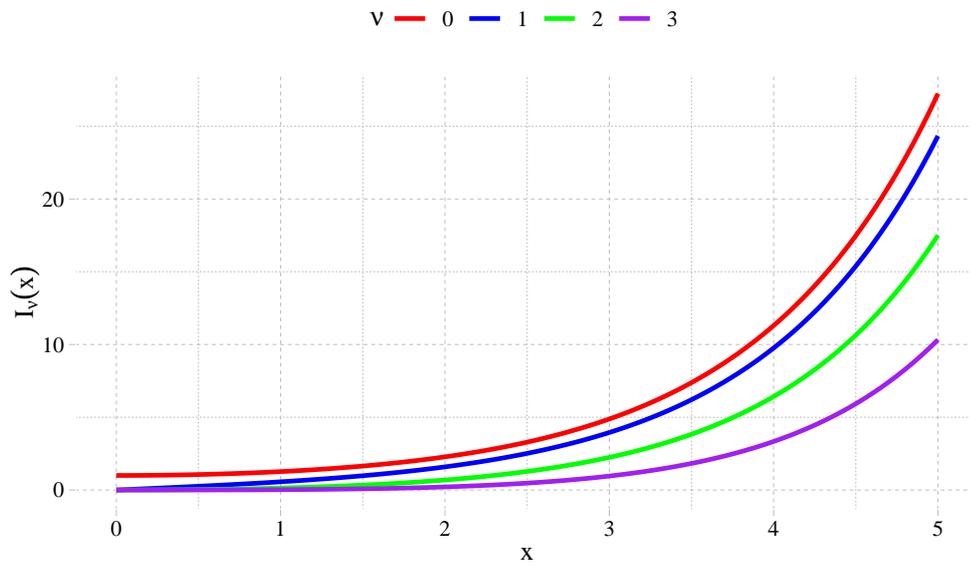
$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} e^{\kappa \cos \theta} d\theta = \sum_{n=0}^{\infty} \left(\frac{\kappa}{2}\right)^{2n} \left(\frac{1}{n!}\right)^2, \quad \kappa \geq 0.$$

A seguir, a Figura 26, apresenta o gráfico da função de Bessel modificada do primeiro tipo para  $\nu = 0, 1, 2, 3$ .

As relações de recorrência satisfeitas por  $I_\nu(x)$  podem ser desenvolvidas pelas expansões de série. No entanto, é mais conveniente trabalhar a partir das relações de recorrência existente para  $J_\nu(x)$  (ARFKEN; WEBER, 2007).

$$\begin{aligned} I_{-\nu}(x) &= I_\nu(x), \\ I_{\nu-1}(x) + I_{\nu+1} &= 2I'_\nu(x), \\ I_{\nu-1}(x) - I_{\nu+1} &= \frac{2\nu}{x} I_\nu(x), \\ \frac{d}{dx} I_0(x) &= I'_0(x) = I_1(x), \end{aligned}$$

Figura 26 – Função de Bessel modificada de primeira espécie:  $I_\nu(x)$



Fonte: O autor (2025).

ainda temos que,

$$\begin{aligned}
 I_0''(x) = I_1'(x) &= \frac{1}{2}[I_0(x) + I_2(x)] \\
 &= I_2(x) + \frac{I_1(x)}{x} \\
 &= I_0(x) - \frac{I_1(x)}{x}.
 \end{aligned}$$

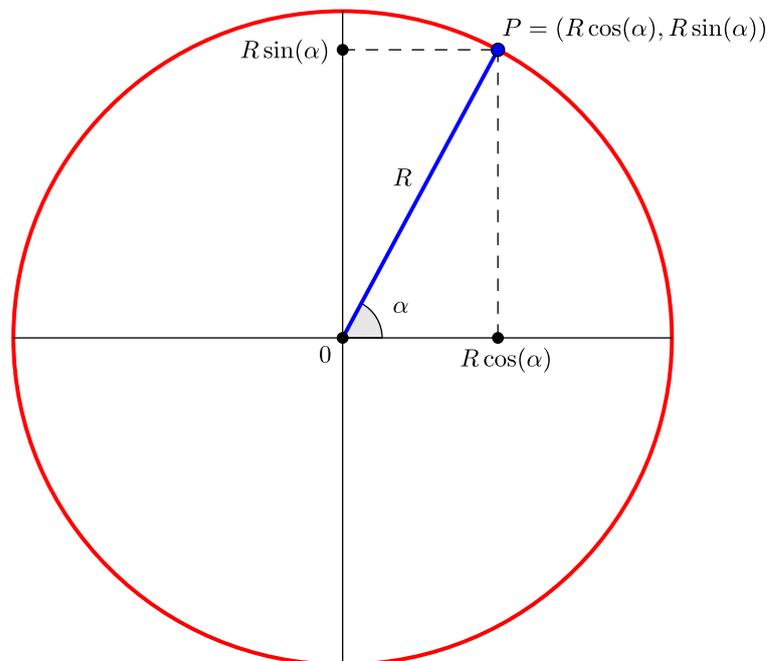
## APÊNDICE B – ESTATÍSTICA CIRCULAR

Existem várias áreas em que os dados são naturalmente modelados de forma circular, como Biologia, Meteorologia, Medicina, Engenharia, Geologia entre outras. Nesses campos, há um interesse significativo em medidas que se apresentam na forma de ângulos, o que é abordado pela Estatística com Dados Circulares. Portanto, é essencial desenvolver uma teoria apropriada para lidar com esses tipos de dados.

### B.1 ANÁLISE DE DADOS CIRCULARES

A estatística circular, trata de técnicas e métodos estatísticos destinados a análise de dados circulares, tendo como estudo as populações ou comunidades no espaço ou no tempo. Podemos representar as grandezas de espaço e tempo em uma circunferência de raio unitário centrada na origem, tanto por meio de pontos quanto de vetores unitários entre a origem e esses pontos. Dessa forma, os dados circulares podem ser representados como ângulos ou pontos na circunferência de um círculo unitário, como podemos observar na Figura 27.

Figura 27 – Relação entre coordenadas cartesianas e polares



Fonte: Adaptada de: (RAO; SENGUPTA, 2001)

A Figura 27, mostra a relação entre as coordenadas cartesianas e polares. Se o ponto  $P$  se encontra na circunferência do círculo unitário, a conversão entre coordenadas cartesianas e

polares é simplesmente

$$(1, \alpha) \Leftrightarrow (x = \cos \alpha, y = \sin \alpha).$$

Sejam  $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$  um conjunto de observações na qual os ângulos são dados em coordenadas polares, e considere a transformação polar para cartesiana em cada observação:

$$(\cos \alpha_i, \sin \alpha_i), \quad i = 1, \dots, n.$$

O vetor resultante dos  $n$  vetores unitários é obtido da seguinte forma:

$$\mathbf{R} = \left( \sum_{i=1}^n \cos \alpha_i, \sum_{i=1}^n \sin \alpha_i \right) = (C, S)$$

onde, o tamanho do vetor resultante fornece informações sobre o grau de concentração dos ângulos observados e pode ser determinado por:

$$R = \|\mathbf{R}\| = \sqrt{C^2 + S^2}.$$

Propõe-se que a direção média circular, que será denotada por  $\bar{\mu}$ , seja dada pela direção do vetor resultante  $R$ . Podemos definir a direção do vetor resultante pelas seguintes equações:

$$\cos \bar{\mu} = \frac{C}{R} \text{ e } \sin \bar{\mu} = \frac{S}{R}.$$

Portanto, o cálculo de  $\bar{\mu}$  é dado por:

$$\bar{\mu} = \begin{cases} \arctan(S/C), & \text{se } C > 0; \\ \arctan(S/C) + \pi, & \text{se } C < 0, S \geq 0; \\ \arctan(S/C) - \pi, & \text{se } C < 0, S < 0; \\ \pi/2, & \text{se } C = 0, S > 0; \\ -\pi/2, & \text{se } C = 0, S < 0; \end{cases} \quad (\text{B.1})$$

A direção média circular  $\bar{\mu}$  é o centro do conjunto de dados e não depende do sentido de rotação ou da escolha da origem. Considerando  $n$  o número de observações, por conveniência, costuma-se definir o comprimento médio do vetor resultante por,  $\bar{R} = R/n$ , o qual pode variar no intervalo  $[0, 1]$ . Deste modo, o valor de  $\bar{\mu}$  pode ser encontrado como soluções das equações:

$$\bar{C} = \bar{R} \cos \bar{\mu} \text{ e } \bar{S} = \bar{R} \sin \bar{\mu} \quad (\text{B.2})$$

onde,

$$\bar{S} = \frac{S}{n} \text{ e } \bar{C} = \frac{C}{n}. \quad (\text{B.3})$$

De acordo com Rao e SenGupta (2001), Arthur, Markus e Graeme (2013), das expressões em (B.2) e (B.3), obtemos:

$$\frac{1}{n} \sum_{i=1}^n \cos(\alpha_i - \bar{\mu}) = \bar{R} \quad \text{e} \quad \frac{1}{n} \sum_{i=1}^n \sin(\alpha_i - \bar{\mu}) = 0. \quad (\text{B.4})$$

Portanto, se as direções  $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$  estão muito concentradas em uma certa parte da circunferência,  $R$  será próximo de 1. Por outro lado, se as observações estiverem dispersas,  $R$  será próximo de 0. Logo,  $R$  é uma medida de concentração do conjunto de dados. Observe que, para qualquer conjunto de dados da forma  $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n, \alpha_1 + \pi, \alpha_2 + \pi, \alpha_3 + \pi, \dots, \alpha_n + \pi$ ,  $R = 0$ . Assim, se  $\bar{R}$  é aproximadamente igual a zero, não necessariamente as observações estão dispersas ao redor da circunferência. Ainda,  $\bar{R}$  é invariante sob rotação.

Na maior parte dos casos, o tamanho do vetor médio resultante,  $\bar{R}$ , é a medida de concentração mais importante. Entretanto, por analogia com a medida para dados lineares, é útil considerar medidas de dispersão circulares. Uma dessas medida é a variância circular amostral, definida como:

$$V = 1 - \bar{R}.$$

## B.2 DISTRIBUIÇÕES DE PROBABILIDADE CIRCULAR

Uma distribuição circular é uma distribuição de probabilidade em que toda a probabilidade está concentrada na circunferência de um círculo unitário. Uma vez que cada ponto na circunferência representa uma direção, essa distribuição é uma maneira de atribuir probabilidades a diferentes direções, constituindo assim uma distribuição direcional. O intervalo de uma VAC,  $\Theta$ , medida em radianos, pode ser tomada no intervalo  $[0, 2\pi)$  ou no intervalo  $(-\pi, \pi]$ . As distribuições circulares são essencialmente de dois tipos: elas podem ser discretas, ou seja, atribuindo massas de probabilidade apenas a uma contagem de direções, ou pode ser absolutamente contínua, (com relação ao Medida de Lebesgue na circunferência). No último caso, uma densidade,  $f_{\Theta}$ , existe e tem as seguintes propriedades básicas:

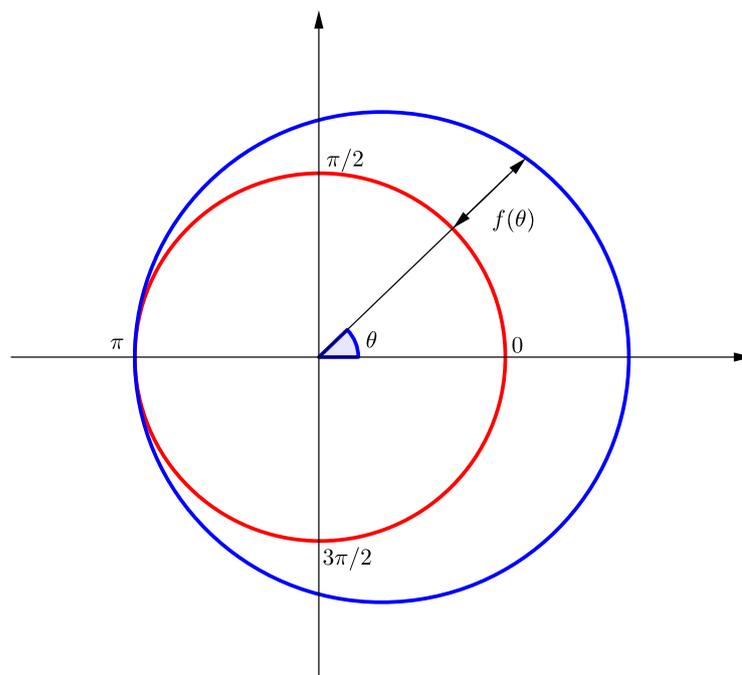
(i)  $f_{\Theta}(\theta) \geq 0$ ;

(ii)  $\int_{-\pi}^{\pi} f_{\Theta}(\theta) d\theta = 1$ ;

(iii)  $f_{\Theta}(\theta) = f_{\Theta}(\theta + 2k\pi)$  para qualquer inteiro  $k$ , ou seja,  $f$  é periódica.

A primeira propriedade,  $f_{\Theta}(\theta) \geq 0$ , garante que a densidade de probabilidade seja não negativa para todos os valores de  $\theta$ . A segunda,  $\int_{-\pi}^{\pi} f_{\Theta}(\theta) d\theta = 1$ , garante que a área sob a curva de densidade seja igual a 1, isto é, probabilidade total igual a 1. A terceira propriedade,  $f_{\Theta}(\theta) = f_{\Theta}(\theta + 2k\pi)$  para qualquer inteiro  $k$ , significa que a densidade de probabilidade é periódica com período  $2\pi$ . Isso indica que a densidade de probabilidade é a mesma para direções diferentes na circunferência, apenas diferindo por múltiplos inteiros de  $2\pi$ . A Figura 28 mostra a representação de uma distribuição circular contínua.

Figura 28 – Representação de uma distribuição circular contínua



Fonte: Adaptado de: (RAO; SENGUPTA, 2001)

Como na linha real, tal distribuição também pode ser descrita por meio de sua função característica. No entanto, como  $\Theta$  é uma VAC periódica tendo a mesma distribuição que  $(\theta + 2\pi)$ , a função característica de tal VAC possui a propriedade:

$$\begin{aligned} \varphi_{\Theta}(t) &\stackrel{\text{def}}{=} E(e^{it\theta}) = E(e^{it(\theta+2\pi)}) \\ &= e^{it2\pi} \cdot \varphi_{\Theta}(t). \end{aligned}$$

Portanto,  $\varphi_{\Theta}(t) = 0$  ou  $e^{it2\pi} = 1$ , ou seja,  $t$  deve ser um número inteiro (RAO; SENGUPTA, 2001). Assim, para VAC's, a função característica precisa ser definida apenas para valores inteiros. O valor da função característica em um inteiro  $p$  também é chamado o  $p$ -ésimo momento trigonométrico de  $\Theta$ . Se  $F_{\Theta}$  denota a CDF da VAC, pode-se escrever: O valor da

função característica em um inteiro  $p$  também é chamado de  $p$ -ésimo momento trigonométrico de  $\Theta$ . Se  $F_\Theta$  denota a função de distribuição acumulada da VAC, pode-se escrever:

$$\begin{aligned}\varphi_\Theta(p) = E(e^{ip\theta}) &= E[\cos(p\theta) + i \sin(p\theta)] \\ &= \int_{-\pi}^{\pi} e^{ip\theta} dF_\Theta(\theta) = \rho_p e^{i\mu_p}, \quad p = 0, \pm 1, \pm 2, \dots,\end{aligned}$$

onde  $\rho_p = |\varphi_\Theta(p)|$  é a magnitude do momento trigonométrico de ordem  $p$ , e  $\mu_p = \arg(\varphi_\Theta(p))$  é seu argumento.

A função característica assume valor 1 no ponto 0, portanto,  $\varphi_0 = 1$ , e

$$|\varphi_p| = \rho_p = \|E(e^{ip\theta})\| \leq E(\|e^{ip\theta}\|) = 1, \quad (\text{B.5})$$

ou seja, a função característica é limitada para todo  $t \in \mathbb{R}$ .

Além disso  $\varphi_p^* = \varphi_{-p}$ , onde,  $\varphi_p^*$  é o conjugado complexo de  $\varphi_p$ . Também podemos ver estes momentos trigonométricos em termos de:

$$\alpha_p = E(\cos p\theta), \quad \beta_p = E(\sin p\theta), \quad p = \pm 0, \pm 1, \pm 2, \dots \quad (\text{B.6})$$

Os momentos em (B.6) estão relacionados pelas equações,

$$\rho_p = \sqrt{\alpha_p^2 + \beta_p^2} \quad \text{e} \quad \mu_p = \arctan\left(\frac{\beta_p}{\alpha_p}\right).$$

Em particular, considere o primeiro momento trigonométrico, dado por,

$$\begin{aligned}\varphi_1 &= \alpha_1 + i\beta_1 \\ &= E(\cos \theta) + iE(\sin \theta) \\ &= \rho_1 \cdot e^{i\mu_1},\end{aligned}$$

que desempenha o papel mais proeminente. Denotando o comprimento  $\rho_1$ , que encontra-se entre 0 e 1 (ver equação (B.5)), simplesmente como  $\rho$  e a direção média  $\mu_1$  simplesmente como  $\mu$ , ou seja,

$$\rho_1 = \rho \quad \text{e} \quad \mu_1 = \mu.$$

Quando o comprimento de  $\rho$  for diferente de zero, a direção  $\mu$ , deste primeiro momento trigonométrico  $\varphi_1$  são usados para fornecer medidas teóricas ou populacionais de concentração e direção média de  $\Theta$ , respectivamente. Observa-se que, quanto maior for o  $\rho$ , ou seja, quanto mais próximo estiver de 1, maior será a concentração no sentido da direção média  $\mu$ .

Considerando  $\mathcal{Z} = e^{i\theta} = \cos \theta + i \sin \theta$  (relação de Euler) como uma variável aleatória, estes são os chamados de momentos brutos de  $\mathcal{Z}$  desde que,

$$\begin{aligned}\varphi_{\Theta}(p) &= E[\mathcal{Z}^p] = E[\mathcal{Z}^{ip\theta}] = E[\cos p\theta + i \sin p\theta] \\ &= E[\cos p\theta] + iE[\sin p\theta] \\ &= \alpha_p + i\beta_p.\end{aligned}$$

Por outro lado, aplicando a expansão binomial a  $\mathcal{Z}$ , obtêm-se:

$$\alpha_p + i\beta_p = \sum_{j=0}^p \binom{p}{j} i^{p-j} E[\cos^j \theta \sin^{(p-j)} \theta],$$

proporcionando uma relação entre os momentos trigonométricos de  $\Theta$  com os momentos usuais obtidos por meio do  $\cos \theta$  e do  $\sin \theta$ .

Por conveniência, a partir de agora, escreveremos  $\varphi_{\Theta}(p)$  simplesmente como  $\varphi_p$ . Estes momentos trigonométricos de  $\Theta$  são iguais aos coeficientes de Fourier na expansão da série de Fourier da função densidade de probabilidade  $f_{\Theta}$ . Para Rao e SenGupta (2001), se a função densidade de probabilidade circular é integrável em  $(-\pi, \pi]$ , é possível recuperá-la a partir de seus coeficientes de Fourier, por meio da expansão:

$$f_{\Theta}(\theta) = \frac{1}{2\pi} \sum_{p=-\infty}^{\infty} \varphi_p e^{-ip\theta} = \frac{1}{2\pi} \left[ 1 + 2 \sum_{p=1}^{\infty} (\alpha_p \cos p\theta + \beta_p \sin p\theta) \right],$$

em que  $\varphi_p = E(e^{ip\theta})$  é o  $p$ -ésimo momento trigonométrico, e  $\alpha_p = \text{Re}(\varphi_p)$ ,  $\beta_p = -\text{Im}(\varphi_p)$ .

Essa expansão é análoga à fórmula de inversão da transformada de Fourier para variáveis reais. O  $p$ -ésimo momento trigonométrico central também pode ser definido por:

$$E(e^{ip(\theta-\mu)}) = \bar{\alpha}_p + i\bar{\beta}_p,$$

onde  $\mu$  é a direção média da variável circular.

Dizemos que a VAC  $\Theta$  é simétrica em torno da direção  $\mu$  se sua densidade satisfaz:

$$f_{\Theta}(\mu + \theta) = f_{\Theta}(\mu - \theta), \quad \forall \theta \in [-\pi, \pi].$$

### B.3 DISTRIBUIÇÕES WRAPPED

Na estatística circular, vários modelos probabilísticos são derivados a partir de transformações de modelos probabilísticos lineares, ou seja, dada uma distribuição sobre a reta, pode-se

embrulhá-la em torno da circunferência do círculo unitário, obtendo assim, uma distribuição de probabilidade angular.

Seja  $X$  uma variável aleatória sobre a reta, com função de densidade de probabilidade  $g(x)$  e, função de distribuição  $G(x)$ . Neste caso, pode-se obter uma variável aleatória circular  $\Theta$  definindo,

$$\Theta \equiv X \pmod{2\pi}$$

Considerando o círculo como o conjunto de números complexos de módulo 1, a função de embrulhamento  $X \mapsto \Theta$  pode ser escrita por:

$$X \mapsto e^{2\pi i X}.$$

Portanto, a função de densidade de probabilidade  $f_{\Theta}(\theta)$  da variável aleatória  $\Theta$  é obtida envolvendo ou embrulhando  $g(x)$  em torno da circunferência do círculo de raio unitário, ou seja,

$$f_{\Theta}(\theta) = \sum_{k=-\infty}^{\infty} g(\theta + 2k\pi), \quad -\pi \leq \theta \leq \pi,$$

que é uma soma periódica de período  $2\pi$ , com função de distribuição correspondente, dada por:

$$F_{\Theta}(\theta) = \sum_{k=-\infty}^{\infty} [G(\theta + 2k\pi) - G(2k\pi)], \quad -\pi \leq \theta \leq \pi.$$

Segue algumas propriedades importantes para essa classe de modelos estatísticos.

1.  $(X + Y)_w = X_w + Y_w$ , onde  $X_w$  e  $Y_w$  são variáveis aleatórias circulares;
2. Se  $X$  tem função característica  $\phi$  então a função característica  $\phi_{\Theta}(p)$  de  $\Theta$  é:

$$\phi_{\Theta}(p) = \phi(p);$$

3. Se  $\phi$  é integrável então  $\Theta$  tem densidade com,  $\sum_{p=1}^{\infty} (\alpha_p^2 + \beta_p^2) < \infty$ , e

$$f_w(\theta) = \sum_{k=-\infty}^{\infty} g(\theta + 2k\pi) = \frac{1}{2\pi} \left( 1 + 2 \sum_{p=1}^{\infty} (\alpha_p \cos(p\theta) + \beta_p \sin(p\theta)) \right).$$

#### B.4 FUNÇÃO CARACTERÍSTICA DA WRAPPED NORMAL

Seja  $X \sim N(\mu, \sigma)$  uma variável aleatória. Então, a função característica de  $X$  é,

$$\phi_X(t) = \exp \left\{ i\mu t - \frac{t^2 \sigma^2}{2} \right\}.$$

Pela propriedade (1) das distribuições Wrapped, segue que a função característica de  $\Theta$  pode ser escrita da seguinte forma:

$$\begin{aligned} \phi(p) &= \exp \left\{ i\mu p - \frac{p^2 \sigma^2}{2} \right\} = \exp\{i\mu p\} \exp \left\{ -\frac{p^2 \sigma^2}{2} \right\} \\ &= \exp\{i\mu p\} \exp \left\{ \left( -\frac{\sigma^2}{2} \right) p^2 \right\} \\ &= e^{i\mu p} \rho^{p^2} \\ &= [\cos(\mu p) + i \operatorname{sen}(\mu p)] \rho^{p^2} \\ &= \rho^{p^2} \cos(\mu p) + i \rho^{p^2} \operatorname{sen}(\mu p). \end{aligned}$$

Portanto,

$$\phi(p) = \alpha_p + i\beta_p,$$

com,

$$\begin{cases} \alpha_p = \rho^{p^2} \cos(\mu p) \\ \beta_p = \rho^{p^2} \operatorname{sen}(\mu p). \end{cases}$$

#### B.5 FUNÇÃO CARACTERÍSTICAS DA WRAPPED CAUCHY

A função característica da distribuição de Cauchy na reta tem a forma:

$$\phi(t) = e^{-\sigma|t| + it\mu}, \quad t \in \mathbb{R}.$$

Logo, a função característica da distribuição Wrapped de Cauchy para  $p \in \mathbb{Z}$  é dada por:

$$\begin{aligned} \phi(p) &= e^{-\sigma|p| + ip\mu} \\ &= e^{-\sigma|p|} e^{ip\mu} \\ &= e^{-\sigma|p|} [\cos(p\mu) + i \operatorname{sen}(p\mu)] \end{aligned}$$

Assim,

$$\phi(p) = e^{-\sigma|p|} [\cos(p\mu)] + e^{-\sigma|p|} [i \operatorname{sen}(p\mu)] = \alpha_p + i\beta_p,$$

em que,

$$\begin{cases} \alpha_p = e^{-\sigma|p|} \cos(p\mu) \\ \beta_p = e^{-\sigma|p|} \operatorname{sen}(p\mu). \end{cases}$$

### B.6 IGUALDADE ENTRE AS EXPRESSÕES DA SEÇÃO 3.1.1.4

A igualdade entre as duas expressões na equação (3.9) é obtida ao relacionar a soma infinita da densidade de probabilidade com uma fórmula fechada, utilizando as propriedades das séries geométricas. Inicialmente, temos que:

$$f_{\Theta}(\theta; \mu, \rho) = \frac{1}{2\pi} \left( 1 + 2 \sum_{k=1}^{\infty} \rho^k \cos(k(\theta - \mu)) \right). \quad (\text{B.7})$$

Utilizamos a identidade de Euler para expressar  $\cos(k(\theta - \mu))$  em termos exponenciais, obtemos:

$$\cos(k(\theta - \mu)) = \frac{e^{ik(\theta - \mu)} + e^{-ik(\theta - \mu)}}{2}. \quad (\text{B.8})$$

Agora, substituindo a expressão em B.8 na soma em B.7, temos:

$$\sum_{k=1}^{\infty} \rho^k \cos(k(\theta - \mu)) = \sum_{k=1}^{\infty} \rho^k \frac{e^{ik(\theta - \mu)} + e^{-ik(\theta - \mu)}}{2}. \quad (\text{B.9})$$

Separando os termos de  $e^{ik(\theta - \mu)}$  e  $e^{-ik(\theta - \mu)}$ , obtemos:

$$\sum_{k=1}^{\infty} \rho^k \cos(k(\theta - \mu)) = \frac{1}{2} \left( \sum_{k=1}^{\infty} (\rho e^{i(\theta - \mu)})^k + \sum_{k=1}^{\infty} (\rho e^{-i(\theta - \mu)})^k \right).$$

Por definição, a soma de uma série geométrica infinita é dada por:

$$\sum_{k=1}^{\infty} a^k = \frac{a}{1 - a}, \quad \text{para } |a| < 1.$$

Para a densidade em B.7, os termos da soma incluem potências de números complexos na forma  $a = \rho e^{i(\theta - \mu)}$  e  $a = \rho e^{-i(\theta - \mu)}$ . Aplicando a fórmula da série geométrica para esses dois casos, obtemos:

$$\sum_{k=1}^{\infty} (\rho e^{i(\theta - \mu)})^k = \frac{\rho e^{i(\theta - \mu)}}{1 - \rho e^{i(\theta - \mu)}}, \quad \sum_{k=1}^{\infty} (\rho e^{-i(\theta - \mu)})^k = \frac{\rho e^{-i(\theta - \mu)}}{1 - \rho e^{-i(\theta - \mu)}}. \quad (\text{B.10})$$

Somando as expressões em B.10 e dividindo por 2, obtemos:

$$\sum_{k=1}^{\infty} \rho^k \cos(k(\theta - \mu)) = \frac{1}{2} \left( \frac{\rho e^{i(\theta - \mu)}}{1 - \rho e^{i(\theta - \mu)}} + \frac{\rho e^{-i(\theta - \mu)}}{1 - \rho e^{-i(\theta - \mu)}} \right). \quad (\text{B.11})$$

Simplificação a expressão em B.11 para obter a fórmula fechada, temos:

$$\frac{\rho e^{i(\theta-\mu)}}{1 - \rho e^{i(\theta-\mu)}} + \frac{\rho e^{-i(\theta-\mu)}}{1 - \rho e^{-i(\theta-\mu)}} = \frac{\rho \left( e^{i(\theta-\mu)}(1 - \rho e^{-i(\theta-\mu)}) + e^{-i(\theta-\mu)}(1 - \rho e^{i(\theta-\mu)}) \right)}{(1 - \rho e^{i(\theta-\mu)})(1 - \rho e^{-i(\theta-\mu)})}. \quad (\text{B.12})$$

Expandindo os termos do numerador da igualdade em B.12.

$$\rho \left( e^{i(\theta-\mu)} + e^{-i(\theta-\mu)} - \rho(e^{i(\theta-\mu)} e^{-i(\theta-\mu)} + e^{-i(\theta-\mu)} e^{i(\theta-\mu)}) \right).$$

Agora, usando  $e^{i(\theta-\mu)} + e^{-i(\theta-\mu)} = 2 \cos(\theta - \mu)$  e  $e^{i(\theta-\mu)} e^{-i(\theta-\mu)} = 1$ , obtemos:

$$\sum_{k=1}^{\infty} \rho^k \cos(k(\theta - \mu)) = \frac{\rho(2 \cos(\theta - \mu) - 2\rho)}{(1 - \rho e^{i(\theta-\mu)})(1 - \rho e^{-i(\theta-\mu)})} = \frac{2\rho \cos(\theta - \mu) - 2\rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)}. \quad (\text{B.13})$$

Assim, usando identidades trigonométricas e substituindo na fórmula da densidade, obtemos:

$$f_{\Theta}(\theta; \mu, \rho) = \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)}.$$