



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

RAMON LIMA DOS SANTOS

ENSAIOS SOBRE TEORIA ASSINTÓTICA EM
RETORNOS SAR SEGUINDO A
DISTRIBUIÇÃO GAMA GENERALIZADA

Recife
2017

RAMON LIMA DOS SANTOS

ENSAIOS SOBRE TEORIA ASSINTÓTICA EM
RETORNOS SAR SEGUINDO A DISTRIBUIÇÃO
GAMA GENERALIZADA

Este trabalho foi apresentado à Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de **Mestre em Estatística**.

ORIENTADOR(A): **Prof. Dr. Abraão David Costa do Nascimento**

CORIENTADOR(A): **Prof. Dr. Alejandro César Frery Orgambide**

RECIFE

2017

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

S237e Santos, Ramon Lima dos
Ensaio sobre teoria assintótica em retornos SAR seguindo a distribuição
Gama generalizada / Ramon Lima dos Santos. – 2017.
81 f.: il., fig., tab.

Orientador: Abraão David Costa do Nascimento.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN,
Estatística, Recife, 2017.
Inclui referências.

1. Estatística. 2. Correção de viés. I. Nascimento, Abraão David Costa do
(orientador). II. Título.

310

CDD (23. ed.)

UFPE- MEI 2017-65

RAMON LIMA DOS SANTOS

**ENSAIOS SOBRE TEORIA ASSINTÓTICA EM RETORNOS SAR SEGUINDO A
DISTRIBUIÇÃO GAMA GENERALIZADA**

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

Aprovada em: 17 de fevereiro de 2017.

BANCA EXAMINADORA

Prof. Abraão David Costa do Nascimento
UFPE

Prof. Renato José de Sobral Cintra
UFPE

Prof. Marcelo Rodrigo Portela Ferreira
UFPB

Dedico esse trabalho primeiramente à Deus.
Como também aos meus pais, Fernando e
Josineide, à minha noiva, Flávia Melina, ao
meu irmão, Renan Lima e ao meu sobrinho,
Luís Augusto.

Agradecimentos

À Deus, amigo incondicional, pelas oportunidades que tem me dado e pela força nas circunstâncias adversas.

À minha noiva, Flávia Melina, pelo amor, companheirismo, incentivo e força nesta reta final do mestrado, a qual peço sinceras desculpas pelos meus períodos ausentes. Sua compreensão, apoio e momentos de descontração foram fundamentais para o desenvolvimento desta dissertação. Sou profundamente grato.

Aos meus pais, Fernando e Josineide, pelo carinho, educação, incentivo e investimento que me deram ao longo destes anos.

À Marcelina Gonzaga, por todo apoio e palavras de ânimo e perseverança.

Agradeço aos meus orientadores, Abraão D. C. Nascimento e Alejandro C. Frery, pela dedicação, compreensão e paciência. Por contribuírem efetivamente com suas experiências na evolução desta dissertação. Pelos conselhos dados em muitas conversas ao fim das reuniões.

Aos meus amigos de longa estrada, como também aos amigos e companheiros de graduação e mestrado. Em especial, aos integrantes da “República Paraibana”: Jodavid, Camila, Pedro, Luana e Andreza.

À Valéria Bittencourt, secretária do programa, por sua competência, dedicação, compreensão e atenção para com os alunos.

À todos os professores do Departamento de Estatística da UFPE. Especialmente aqueles que participaram da minha formação durante o mestrado: Francisco Cribari, Alex Dias, Audrey Helen Cysneiros, Abraão Nascimento e Gauss Cordeiro.

À todos os professores do Departamento de Estatística da UFPB, por contribuírem intensamente em minha formação acadêmica na graduação.

À CAPES, pelo apoio financeiro. E, finalmente, aos membros da banca examinadora, por suas valiosas sugestões na finalização deste trabalho.

“Você pode encarar um erro como uma besteira a ser esquecida, ou como um resultado que aponta uma nova direção.”

Steve Jobs

Resumo

Os sistemas de radar de abertura sintética (*Synthetic Aperture Radar-SAR*) têm sido apresentados como ferramentas eficientes na resolução de problemas de sensoriamento remoto. Tais sistemas exibem várias vantagens. Em particular, seu funcionamento independe do horário do dia e/ou das condições atmosféricas, como também eles podem fornecer imagens em alta resolução espacial. Entretanto, as imagens SAR são contaminadas por um tipo de interferência denominada ruído *speckle*, dificultando o reconhecimento de padrões em tais imagens por análise visual e/ou pelo uso de métodos clássicos de processamento automático. Assim, a proposta de técnicas estatísticas especializadas (incluindo modelagem e métodos inferenciais) se torna uma importante etapa no processamento e na análise de imagens SAR. Recentemente, uma versão da distribuição gama generalizada tem sido aplicada com sucesso para descrever dados SAR em formato de intensidade. Esta dissertação apresenta inicialmente uma discussão sobre estimação por máxima verossimilhança (MV) para os parâmetros do modelo gama generalizado adaptado ao contexto de dados SAR. Adicionalmente, derivamos um teorema que permite encontrar analiticamente a matriz informação de Fisher do modelo gama generalizado. Estimadores em forma fechada são propostos (sendo um deles para o número de *looks*, parâmetro que descreve o efeito do ruído sobre imagens SAR) e um terceiro é definido como solução de uma equação não-linear. Em segundo lugar, propomos um método de estimação melhorado para os parâmetros da distribuição gama generalizada, derivando a expressão do viés de segunda ordem de acordo com a proposta de Cox e Snell [Journal of Royal Statistical Society. Series B (Methodological), vol. 30, no. 2, pp. 248–275, 1968]. Finalmente, objetivando definir contrastes no espaço paramétrico da lei gama generalizada, derivamos seis medidas de divergência (discrepância entre duas medidas de probabilidade) com base na medida de Kullback-Leibler simetrizada e aplicamos os conceitos de testes de hipóteses fundamentados nas propriedades assintóticas da classe de divergências estudada por Salicrú *et al.* [Journal of Multivariate Analysis, vol. 51, pp. 372–391, 1994]. Três dentre as medidas derivadas quantificam o erro em escolher o modelo gama quando um fenômeno é regido

pela distribuição gama generalizada, ou vice-versa. Duas outras medidas calculam o contraste entre dois elementos diferentes a partir da distribuição gama generalizada ou, como caso particular, da distribuição gama. Outra quantifica o erro de escolher o modelo gama generalizado trivariado não-correlacionado quando os dados seguem o modelo Wishart complexo escalonado. Esta última quantidade pode ser utilizada como um detector de redundância para uma região SAR polarimétrica; isto é, um teste de hipótese que informa quando trabalhar com modelos marginais gama generalizado é estatisticamente similar a usar distribuições matriciais. As demais medidas são definidas ou como testes de hipóteses para duas regiões de intensidades SAR ou como medidas de bondade de ajuste entre gama e gama generalizada no mesmo contexto. Para quantificar a eficiência das novas metodologias, estudos de simulação Monte Carlo são realizados bem como vários experimentos com dados SAR reais.

Palavras-chave: Correção de viés. Imagem SAR. Gama generalizada. Wishart. Divergência.

Abstract

Synthetic aperture radar (SAR) systems have been presented as efficient tools for remote sensing. Such systems exhibit several advantages. In particular, its operation is independent of the time of day and/or atmosphere conditions, as well as they can provide images in high spatial resolution. However, SAR images are contaminated by a type of interference called speckle noise, hindering the recognition of patterns in such images by visual analysis and/or by the use of classical automatic processing methods. Thus, the proposal of specialized statistical techniques (including modeling and inferential methods) is an important step in the processing and analysis of SAR images. Recently, a version of the generalized gamma distribution (GG) has been successfully applied to describe SAR data in intensity format. This dissertation presents first a discussion about estimation by maximum likelihood (ML) for the parameters of the GG model tailored to the SAR data context. Additionally, we derive a theorem that allows to find analytically the Fisher information matrix of the GG model. Furthermore, two closed-form estimators are proposed (being one of them for the number of looks, parameter which describes the effect of noise on SAR images) and a third which is defined as a solution of one non-linear equation. Second we propose an improved estimation method for the GG parameters, deriving the second-order bias expression according to the proposal of Cox and Snell [Journal of Royal Statistical Society. Series B (Methodological), vol. 30, no. 2, pp. 24–275, 1968]. Finally, aiming to define contrasts in the GG law parametric space, we derive six divergence measures (discrepancy between two probability measures) based on the Kullback-Leibler measure symmetrized and we apply the concepts of hypothesis testing based on the asymptotic properties from the h - ϕ class studied by Salicrú et al. [Journal of Multivariate Analysis, vol. 51, pp. 37–391, 1994]. Three of among derived measures quantify the error in choosing the GG model when a phenomenon is governed by the GG distribution, or conversely. Two other measures compute the contrast between two different elements modeled by the GG or the Γ distribution. Other measure quantifies the error of choosing the uncorrelated trivariate GG model when the data follow the scaled complex Wishart model.

This latter quantity can be used as a redundancy detector for a polarimetric SAR region; i.e., a hypothesis test that tells us when to employ marginal $G\Gamma$ models is statistically similar to use matrix distributions. The other five previous measures are defined either as hypothesis tests for two SAR intensity regions or as criteria of goodness of fit between Γ and $G\Gamma$ in the same context. To quantify the efficiency of the new methodologies, Monte Carlo simulation studies are performed as well as several experiments with actual SAR data.

Keywords: Bias Correction. SAR Image. $G\Gamma$. Wishart. Divergence.

Lista de Figuras

1.1	Sistema AIRSAR	20
1.2	Baía de São Francisco (EUA)	21
1.3	Recorte da imagem AIRSAR da Baía de São Francisco (EUA).	21
2.1	Diagrama ilustrando casos particulares do modelo $G\Gamma$ para dados de intensidade SAR.	25
2.2	Curvas de densidade do modelo $G\Gamma$ para $[\theta, L, \mu] = [\bullet, 0.1, 3]$	26
2.3	Imagem sintética para $L \in \{1, 3, 20\}$ e $\theta \in \{1, 2, 5\}$	27
2.4	Recorte da imagem AIRSAR de São Francisco com <i>pixels</i> destacados.	31
4.1	Estudo das propriedades assintóticas dos vieses de $\hat{\mu}(\theta)$ e $\hat{L}(\theta)$ sob $\mu = 1$, $L = 1$ e $ \theta \in \{0.8, 1.0, 4.0\}$	50
4.2	Densidades empíricas (\bullet) vs. ajustadas para distribuições $G\Gamma$ (curvas sólidas pretas), $G\Gamma$ especificado $L = 4$ (curvas tracejadas pretas) e Γ (curvas sólidas cinzas).	53
4.3	p -valores para os testes da razão entre verossimilhanças sob $\mathcal{H}_0: G\Gamma = \Gamma$ e $N = 49, 81, 121$, usando janelamentos $7 \times 7, 9 \times 9$ e 11×11 , respectivamente.	56
4.4	Estimativas de L sob o Modelo Γ para $N = 49, 81, 121$, usando janelamentos $7 \times 7, 9 \times 9$ e 11×11 , respectivamente.	58
4.5	Estimativas de L sob o Modelo $G\Gamma$ para $N = 49, 81, 121$, usando janelamentos $7 \times 7, 9 \times 9$ e 11×11 , respectivamente.	59
4.6	Estimativas de L para nossa proposta e proposta por Anfinsen, $N = 9, 25, 49$, usando janelamentos $7 \times 7, 9 \times 9$ e 11×11 , respectivamente.	60
4.7	p -valores para os testes da razão entre verossimilhanças sob $\mathcal{H}_0: G\Gamma = G\Gamma_{cor}$ e $N = 9, 49$, para janelamentos 3×3 e 7×7 , respectivamente.	63
4.8	Imagem sintética $G\Gamma$ para $(L, \text{sign}(\theta)) \in \{(1, 1), (8, 1), (1, -1), (8, -1)\}$ e $\theta \in \{3, 5, 8, 15\}$	64

4.9	Curvas de poder do teste para a estatística proposta.	66
4.10	Mapa das estimativas de θ para ajustes Γ e $G\Gamma$ em janela 10×10	67
4.11	Mapa dos valores de S_{DZY} , S_{DYZ} e S_{dZZ} com base em uma janela 10×10	68
4.12	Divergência KL entre os modelos $G\Gamma$ e \mathcal{W}	69
4.13	Divergência KL entre os modelos Γ e \mathcal{W}	70
4.14	p -valor do teste de hipótese baseado na divergência KL entre os modelos $G\Gamma$ e \mathcal{W}	71
4.15	p -valor do teste de hipótese baseado na divergência KL entre os modelos Γ e \mathcal{W}	72

Lista de Tabelas

2.1	Tabela <i>Pixel-M</i> baseada em recorte da imagem AIRSAR de São Francisco.	32
4.1	Performance dos EMVs para o parâmetro θ do modelo GF	49
4.2	Estimativas MV para L e seus respectivos EQMs para dados SAR sintéticos	51
4.3	Estatísticas descritivas para os dados SAR coletados	52
4.4	EMVs e seus EPs estimados para os dados SAR de intensidade	54
4.5	Estatísticas RV e seus p-valor sob $\mathcal{H}_0: GF = \Gamma$	55
4.6	Medidas de bondade de ajuste para os modelos GF , $GF \mathcal{L} = 4$ e Γ nos canais HH, HV e VV	57
4.7	Estimativas MV para L e seus respectivos EQMs para dados SAR reais . .	57
4.8	Estimativas MV para L corrigidas e não corrigidas e seus respectivos EQMs para dados SAR sintéticos	61
4.9	Estimativas MV para θ corrigidas e não corrigidas e seus respectivos EQMs para dados SAR sintéticos	62
4.10	Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ em que $ \theta = 1, \mu = 1, L \in \{1, 3, 4, 8\}$ e tamanho amostral N . .	65
4.11	Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ em que $ \theta < 1, \mu = 1, L \in \{1, 3, 4, 8\}$ e tamanho amostral N . .	73
4.12	Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ em que $ \theta > 1, \mu = 1, L \in \{1, 3, 4, 8\}$ e tamanho amostral N . .	74
4.13	Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ e tamanho amostral N para distribuições GF baseada nos dados	74
4.14	Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ e tamanho amostral N para distribuições $GF L = 4$ baseada nos dados	75
4.15	Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ e tamanho amostral N para distribuições Γ baseada nos dados .	75

Sumário

1	Introdução	16
1.1	Motivação	16
1.2	Contribuições desta dissertação	19
1.3	Plataforma computacional e dados de imagem real utilizados	20
1.4	Organização da dissertação	22
2	Modelo $G\Gamma$ para Intensidades SAR: Revisão e contribuições	23
2.1	Modelo $G\Gamma$	24
2.2	Inferência baseada em verossimilhança para o modelo $G\Gamma$	26
2.3	Refinamentos para inferência na distribuição $G\Gamma$: Estimador semiparamétrico para o NEL	30
2.4	Estimador corrigido segundo Cox-Snell para os parâmetros do modelo $G\Gamma$.	32
2.4.1	Viés para modelo $G\Gamma$ escalonado	34
2.4.2	Viés para modelo $G\Gamma$ não escalonado	38
3	Teoria da Informação Estatística: Divergência Estocástica para os modelos $G\Gamma$ e \mathcal{W}^C	41
3.1	Novas medidas de contraste com base na distribuição $G\Gamma$	43
3.2	Nova medida de contraste entre distribuições \mathcal{W}_3^C e $G\Gamma$ (Γ)	45
4	Resultados	47
4.1	Estimação por MV para os parâmetros do modelo $G\Gamma$ e estimador semi paramétrico para o número de <i>looks</i>	48
4.1.1	Desempenho em imagens sintéticas	48
4.1.2	Aplicação com dados SAR de intensidade	52
4.2	Estimação melhorada para o modelo $G\Gamma$: Vieses de segunda ordem de Cox e Snell	60

4.2.1	Estudo com dados sintéticos	60
4.2.2	Aplicação com dados SAR de intensidade	62
4.3	Medidas de divergência com base nas distribuições $G\Gamma$ e WCE	63
4.3.1	Estudo sintético	63
4.3.2	Aplicação com dados SAR de intensidade	67
5	Considerações finais	76
	Referências	78

1.1 Motivação

O geosensoriamento de um meio físico requer a utilização de técnicas sofisticadas. Dentre elas, destaca-se o sensoriamento remoto que objetiva capturar, processar e analisar informações de fenômenos naturais, a partir da interação de um alvo com radiação eletromagnética (JENSEN, 2007).

Os sistemas de radar de abertura sintética (*Synthetic Aperture Radar*–SAR) têm sido indicados como ferramentas importantes na resolução de problemas de sensoriamento remoto (OLIVER; QUEGAN, 1998). Estes sistemas possuem várias vantagens (LEE; POTTIER, 2009), incluindo a capacidade de operar independentemente da luminosidade e das condições climáticas, assim como fornecer imagens em alta resolução espacial. Por este motivo, imagens SAR têm sido amplamente utilizadas em diversas atividades, tais como: monitoramento ambiental, planejamento urbano, reconhecimento de culturas agrícolas, mapeamento da superfície terrestre e detecção de minas (LI *et al.*, 2010; LI *et al.*, 2011).

Entretanto, devido à utilização de *iluminação coerente*¹ na etapa de aquisição, imagens SAR são fortemente contaminadas por um padrão de interferência chamado *ruído speckle* (LEE; POTTIER, 2009). Este fenômeno dificulta consideravelmente a modelagem, a interpretação e o processamento das imagens resultantes. Assim, o uso de modelos probabilísticos adequados e o desenvolvimento de métodos inferenciais especializados se tornam fundamentais durante o pré-processamento de atributos de dados SAR (NASCIMENTO *et al.*, 2010).

Estes dados podem ser entendidos como o resultado da emissão e captura de pulsos

¹Em função da característica coerente do sensor, a amplitude e a fase do sinal recebido (representado por um número complexo) são resultantes de uma soma vetorial (*coerente*) dos diversos elementos de um alvo natural (FREITAS *et al.*, 2007)

polarizados linearmente nas direções horizontal (“H”) e vertical (“V”) implicando na associação de um vetor complexo, diga-se $(S_{VV}, S_{VH}, S_{HV}, S_{HH}) \in \mathbb{C}^4$, para cada entrada da imagem. Na prática, sob as condições do teorema de reciprocidade (ULABY; ELACHI, 1990), $S_{VH} = S_{HV}$ e um elemento associado ao canal $k \in \{VV, HV, HH\}$ é definido como

$$S_k = |S_k|e^{j\phi_k},$$

em que $\phi_k \in [0, 2\pi]$ é a fase de S_k . Nesta dissertação, tem-se dentre um de seus interesses, propor refinamentos estatísticos para um modelo de I_k , intensidade de S_k , em que $I_k = |S_k|^2$.

Recentemente, Gao (2010) propôs uma pesquisa que descreve quatro abordagens para a modelagem de intensidades em dados SAR, destacando-se uma que é classificada como distribuições empíricas. Neste âmbito, as distribuições log-normal, Weibull e Fisher (que é uma reparametrização da lei \mathcal{G}^0 inicialmente estudada por Frery *et al.* (1997)) têm sido indicadas como modelos importantes. Ainda dentro da abordagem intitulada como distribuições empíricas, vários trabalhos na literatura de dados SAR indicaram a distribuição Γ como uma suposição estocástica para dados *singlelook* e *multilook* pouco texturados (GIERULL; SIKANETA, 2002). No entanto, aplicações a dados de intensidade reais têm indicado o modelo gama generalizado ($G\Gamma$) como mais adequado do que o Γ e outros na descrição de cenários com diferentes texturas; tais como, regiões de pasto e oceano (áreas pouco texturadas), florestas (áreas moderadamente texturadas) e regiões urbanas (áreas intensamente texturadas) (LI *et al.*, 2010). Em particular, Li *et al.* (2011) propuseram uma versão da distribuição $G\Gamma$ como um modelo para descrever intensidades SAR, bem como um método de estimação baseado em cumulantes de segunda ordem.

Uma importante etapa de pós-processamento é a definição de um método de estimação eficiente para os parâmetros que indexam as distribuições probabilísticas de interesse. O método de máxima verossimilhança (MV) é frequentemente utilizado em aplicações SAR por suas propriedades assintóticas, como centralidade e normalidade em grandes amostras. No caso em que a distribuição $G\Gamma$ está parametrizada segundo a versão de Stacy (STACY, 1962), Hager e Bain (1970) e Gomes *et al.* (2008) discutiram o processo de estimação por MV. Entretanto, o estimador de máxima verossimilhança (EMV) é frequentemente tendencioso para amostras finitas, tendo viés $O(n^{-1})$ em que n é o tamanho amostral (CORDEIRO, 2011). Tal viés não é considerado em situações práticas, pois seu valor é desprezível em relação ao erro padrão, que está em $O(n^{-1/2})$. No entanto, tais vieses podem ser substanciais em situações em que são utilizados tamanhos de amostra pequenos ou moderados. Convém ressaltar que comumente métodos de processamento de imagens utilizam janelas com número de pixels pequeno; como, por exemplo, em métodos de filtragem de dados SAR (TORRES *et al.*, 2014). Assim, a proposta de expressões analíticas para o viés do EMV são requeridas.

Cox e Snell (1968) propuseram uma fórmula geral para o viés de segunda ordem do EMV. Desde então, vários trabalhos empregaram esta fórmula como um meio de obter

estimadores melhorados para parâmetros de várias distribuições (CORDEIRO, 2011). Vasconcellos *et al.* (2005) aplicaram-na à distribuição \mathcal{G}^0 univariada, que é frequentemente usada para modelar imagens SAR.

Nesta dissertação propõem-se versões de estimadores melhorados para os parâmetros da distribuição $\mathcal{G}\Gamma$ segundo a abordagem de Cox-Snell. Em particular, a literatura de SAR reserva um interesse especial pelo número de equivalente de *look* (NEL). Esta quantidade pode ser definida como um parâmetro do modelo $\mathcal{G}\Gamma$. Como uma segunda proposta desta dissertação, apresenta-se um estimador semi-paramétrico para o NEL. Este último é definido por meio de uma combinação entre estimadores para NEL a partir do modelo $\mathcal{G}\Gamma$ e a decomposição espectral da matriz de coerência².

Adicionalmente à proposta de estimação melhorada como *input* para métodos de pós-processamento de imagem, a definição de medidas de discrepâncias eficientes é uma etapa crucial de pré-processamento. Assim, o uso de medidas da Teoria da Informação tem se mostrado promissor na análise de dados de imagens SAR.

Goudail *et al.* (2004) aplicaram medidas estocásticas para caracterizar o desempenho de algoritmos de detecção de alvos e de segmentação no processamento de imagens SAR polarimétricas (PolSAR). Nesse estudo, as distâncias de Kullback-Leibler e Bhattacharyya foram empregadas como ferramentas para quantificar a dissimilaridade entre as distribuições Gaussiana complexa circulares. No entanto, as propriedades estatísticas das medidas usadas não foram explicitamente analisadas.

Em contrapartida, recentemente alguns trabalhos foram propostos observando aquelas propriedades, como Nascimento *et al.* (2010), Nascimento *et al.* (2014), Frery *et al.* (2013b), Frery *et al.* (2014). Com base nas distribuições assintóticas das estatísticas obtidas a partir das distâncias estocásticas, Nascimento *et al.* (2010) propuseram testes de hipótese para verificar a igualdade entre modelos que seguem a distribuição \mathcal{G}_T^0 ; enquanto que Frery *et al.* (2014) propuseram testes para avaliar a equivalência entre modelos Wishart complexo escalonados (WCE).

A divergência de Kullback-Leibler (KL) tem apresentado notável importância na literatura de Teoria da Informação (COVER; THOMAS, 1991), sendo uma das mais antigas medidas de discriminação entre as distribuições de probabilidade. Nos últimos anos, essa ferramenta tem sido aplicada em campos de Inferência Estatística: tais como medidas de bondade de ajuste (JAGER; WELLNER, 2007) e testes de hipóteses (TOMA; LEONI-AUBIN, 2009). Em particular, Eguchi e Copas (2006) propuseram uma nova interpretação do lema de Neyman-Pearson em termos da divergência KL.

Nesta dissertação, medidas de distância KL para o modelo $\mathcal{G}\Gamma$ são propostas e suas propriedades estatísticas são analisadas no contexto de processamento de imagens SAR.

Sabendo que o ruído *speckle* complica a modelagem, a interpretação e o processamento de dados de intensidade SAR, esta dissertação faz contribuições assumindo o uso da dis-

²Matriz introduzida por Cloude (1986) e utilizada na decomposição de alvos incoerentes (CLOUDE; POTTIER, 1996). O uso da matriz de coerência é preferido na literatura devido a interpretabilidade física dos seus elementos.

tribuição $G\Gamma$ como uma alternativa eficiente na descrição de tais dados, comparando-a com as distribuições Γ e Wishart. Então, utiliza-se o método de MV devido suas propriedades assintóticas, para estimar os parâmetros dos três modelos e, através do viés e erro quadrático médio (EQM), avalia-se a precisão dos estimadores. Porém, tais estimadores podem se apresentar bastante viesados quando o tamanho amostral é pequeno ou moderado. Assim, a derivação de expressões que possibilitem reduzir o viés se torna importante para a obtenção de estimadores mais precisos que os não-corrigidos. Desta forma, adicionalmente, apresenta-se expressões de forma fechada para o cálculo do viés de segunda ordem dos estimadores de máxima verossimilhança do modelo $G\Gamma$ a partir da expressão geral dada por Cox e Snell (1968). Ainda, propõe-se medidas de distância KL para a distribuição $G\Gamma$ e testes de hipóteses baseados nestas para verificar a igualdade entre os modelos apresentados. Segue-se detalhamento das contribuições desta dissertação.

1.2 Contribuições desta dissertação

1^a contribuição

Uma reparametrização do modelo $G\Gamma$ para intensidades SAR pode incluir o número de looks (parâmetro que objetiva melhorar a razão sinal-ruído) e um parâmetro de forma (que flexibiliza a distribuição na descrição de diferentes texturas). Neste contexto, propõe-se um teorema que permite encontrar analiticamente a matriz de informação de Fisher do modelo $G\Gamma$. Além disso, dois estimadores em forma fechada são propostos (sendo um deles para o número de looks) e um terceiro que é definido como uma solução de uma equação não-linear. Adicionalmente, com base nos EMVs para o número de looks e na decomposição espectral da matriz de coerência, um estimador semi-paramétrico para o número de *looks* que pode fornecer melhor *performance* que outros existentes na literatura SAR.

2^a contribuição

Proposta de um método de estimação MV para a distribuição $G\Gamma$ por meio da derivação do viés de segunda ordem de acordo com a expressão proposta por Cox e Snell [Journal of the Royal Statistical Society. Series B (Methodological), vol. 30, no. 2, pp. 248–275, 1968];

3^a contribuição

Derivação de seis medidas de divergência (que podem ser entendidas como testes de hipóteses para atributos dos dados SAR de intensidade) a partir da combinação entre a divergência de Kullback-Leibler simetrizada e as propriedades assintóticas da classe de divergências $h-\phi$, estudada por Nascimento *et al.* [IEEE Transactions on Geoscience and Remote Sensing, vol. 48, pp. 373–385, 2010]. Três dentre as medidas a serem derivadas quantificam o erro pela escolha do modelo Γ quando os dados empíricos seguem a distri-

buição $G\Gamma$, ou vice-versa. Por outro lado, mais duas medidas calculam o contraste entre dois elementos diferentes a partir da distribuição $G\Gamma$ ou, como caso particular, da distribuição Γ . Uma última que quantifica o erro pela escolha do modelo $G\Gamma$ trivariado não correlacionado quando os dados seguem o modelo WCE será fornecida.

1.3 Plataforma computacional e dados de imagem real utilizados

Linguagem de programação R

Para produção de gráficos e imagens, utilizou-se a linguagem de programação **R** (CHAMBERS, 2008). Em particular, funções foram criadas para os seguintes fins:

- a. Leitura dos dados de intensidade SAR;
- b. Estimação dos parâmetros;
- c. Correção de viés por meio da expressão de Cox-Snell;
- d. Cálculo de medidas de contraste em dados SAR;
- e. Realização de simulações e experimentos reais.

Sensor AIRSAR (*Airborne Synthetic Aperture Radar-AIRSAR*) e dados utilizados

O sensor aerotransportado AIRSAR foi projetado e desenvolvido pelo laboratório de propulsão à jato (*Jet Propulsion Laboratory-NASA-JPL*) em 1980 e sua primeira versão foi acoplada à aeronave CV-990. Entretanto, esta versão foi absolutamente destruída devido um acidente envolvendo a aeronave em julho de 1985. Após, uma nova versão conhecida por AIRSAR, cuja imagem é exibida na Figura 1.1, foi desenvolvida pelo JPL. Esta versão operou em modo polarimétrico sob as bandas P-(0,45 GHz), L-(1,26 GHz) e C-(5,31 GHz) simultaneamente. O sistema AIRSAR concebeu seu primeiro voo em 1987, realizando no mínimo uma atividade por ano a partir de então.



Figura 1.1: Sistema AIRSAR

Como base de dados, esta dissertação utilizou uma imagem da baía de São Francisco (EUA) capturada pelo sensor AIRSAR na banda L e adotando um número de *looks*

quatro. Esta imagem PolSAR é utilizada em muitos artigos, pois apresenta três regiões notavelmente distintas: oceano, floresta e zonas urbanas. Estas regiões possuem diferentes graus de textura bem definidos, submetendo os métodos utilizados a cenários variados. A seguir, a Figura 1.2 (Fonte: http://earth.eo.esa.int/polsarpro/Manuals/021_polarimetric_airborne_sensors.pdf) apresenta a referida imagem com coloração conforme a representação de Pauli³, assim como um mapa especificando as regiões. Por questões de simplificação computacional, utilizou-se três regiões selecionadas em recorte da Figura 1.2(b) conforme apresentado na Figura 1.3.



(a) Mapa de São Francisco

(b) Imagem AIRSAR de São Francisco representada pela configuração de Pauli

Figura 1.2: Baía de São Francisco (EUA)

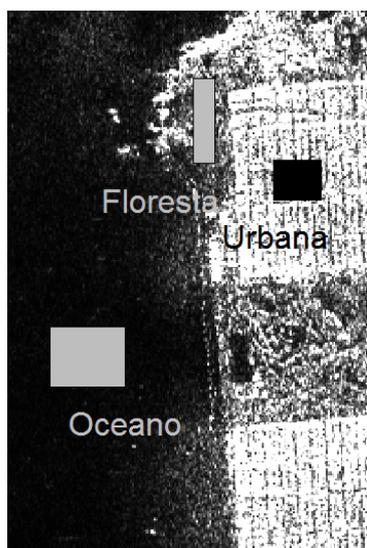


Figura 1.3: Recorte da imagem AIRSAR da Baía de São Francisco (EUA).

³A *representação de Pauli* tem recebido atenção especial por delinear a imagem sobre as combinações HH+VV, HH-VV, HV as relacionando às cores azul, vermelho e verde. Isto possibilita acrescentar coloração a uma imagem PolSAR permitindo destacar algumas de suas características importantes.

1.4 Organização da dissertação

O restante desta dissertação está organizado da seguinte forma:

- O Capítulo 2 apresenta uma discussão sobre estimação por MV para os parâmetros do modelo GF. Uma proposta de estimação melhorada por meio da expressão para correção de viés de Cox-Snell também é proposta. Adicionalmente, apresenta-se um estimador semi-paramétrico para o número de *looks*.

- O Capítulo 3 descreve a metodologia de testes de hipóteses baseados na classe de divergências (h, ϕ) proposta por Salicrú *et al.* (1994). Aqui, apresentam-se medidas de divergência estocástica para os modelos GF e WCE.

- O Capítulo 4 aborda uma discussão numérica sobre as principais propostas desta dissertação. Inicialmente, as propostas citadas são submetidas a dados sintéticos via simulação Monte Carlo. Em seguida, experimentos com dados reais de intensidade SAR são realizados.

- Finalmente, o Capítulo 5 apresenta as considerações finais.

Modelo $G\Gamma$ para Intensidades SAR: Revisão e contribuições

Imagens SAR são obtidas obedecendo a seguinte dinâmica (LEE; POTTIER, 2009): Pulsos polarizados (nas direções horizontal, ‘H’, and vertical, ‘V’) são enviados a uma superfície geográfica e seus retornos são registrados. No contexto univariado, uma imagem é produzida de modo que cada pixel seja representado por um número complexo chamado *canal de polarização*. Este canal pode ser associado a um dentre quatro estados: (i) horizontal-horizontal, ‘HH’, (ii) horizontal-vertical, ‘HV’, (iii) vertical-horizontal, ‘VH’, e (iv) vertical-vertical, ‘VV’. O seu valor para a entrada (m, n) de uma imagem SAR pode ser definido como

$$X(m, n) = X_{\text{Re}}(m, n) + \mathbf{j} X_{\text{Im}}(m, n) = A \exp\{\mathbf{j}\phi\} = \sum_{k=1}^N A_k \exp\{\mathbf{j}\phi_k\}, \quad (2.1)$$

em que N é o número de *scatters*¹ discretos em cada célula de resolução, $\mathbf{j} = \sqrt{-1}$ é a unidade imaginária, e A_k e ϕ_k representam a amplitude e a fase no k -ésimo *scatter*, respectivamente. Adicionalmente, $X_{\text{Re}}(m, n) = \Re[X(m, n)]$ e $X_{\text{Im}}(m, n) = \Im[X(m, n)]$ são as componentes em *fase* e em *quadratura* do canal de polarização (ou sinal SAR complexo retroespalhado) $X(m, n)$, respectivamente, $\Re[\cdot]$ e $\Im[\cdot]$ são as partes real e imaginária de um argumento, respectivamente, e ϕ , A e $I = A^2$ representam a fase, a amplitude e a intensidade de $X(m, n)$, respectivamente. Em seguida, assumimos a distribuição $G\Gamma$ como um modelo para descrever I , o qual estende o modelo Γ para o ruído *speckle* em processos *multilook*.

Obtidas a partir do registro coerente de ecos retornados, imagens SAR são contaminadas por ruído *speckle*. Devido a desvios das propriedades clássicas de (i) aditividade e

¹Elementos dentro de uma célula de resolução (NASCIMENTO, 2012).

(ii) normalidade para $X_{\text{Re}}(m, n)$ e $X_{\text{Im}}(m, n)$, as imagens contaminadas pelo ruído requerem processamento específico.

Na prática, o ruído *speckle* em intensidade de imagens *singlelook* é distribuído exponencialmente com média $\mu > 0$ (LI *et al.*, 2011). Portanto, um procedimento *multilook* (que consiste em computar o valor médio em uma amostra de $L \geq 1$ retornos independentes referentes a *pixels* adjacentes (LEE; POTTIER, 2009)) fornece imagens cujas intensidades dos retornos *multilook* podem ser descritas pela distribuição gama, $Y \sim \Gamma(L, \mu^{-1}L)$ (FRERY *et al.*, 1997), com função densidade de probabilidade (fdp) dada por

$$f_Y(y; L, \mu) = \frac{L^L}{\mu \Gamma(L)} \left(\frac{y}{\mu}\right)^{L-1} \exp\left[-L\frac{y}{\mu}\right] \mathbb{I}_{(0, \infty)}(y), \quad (2.2)$$

em que $\mathbb{I}_A(\cdot)$ é a função indicadora no conjunto A , $L > 0$ representa o número de looks e $\mu > 0$ é a média do retorno. Li *et al.* (2011) propuseram o seguinte lema.

Lema 1. *Seja $Y \sim \Gamma(L, \mu^{-1}L)$, então os seguintes valores esperados são válidos:*

- a) $E\left\{\log\left(\frac{Y}{\mu}\right)\right\} = [\psi^{(0)}(L) - \log(L)];$
- b) $E\left(\frac{Y}{\mu}\right) = 1;$
- c) $E\left\{\left(\frac{Y}{\mu}\right)^\theta\right\} = L^{-\theta} \frac{\Gamma(L+\theta)}{\Gamma(L)},$

em que $E(\cdot)$ representa o valor esperado de uma variável aleatória e $\psi^{(0)}(x)$ é a função digama.

Na prática, quando o parâmetro L (que é um número inteiro) é estimado a partir de dados reais, procura-se obter aquele que torne a fdp Γ mais ajustável a um determinado conjunto de dados. Um estudo detalhado sobre a estimação desta quantidade foi fornecido por Anfinsen *et al.* (2009) e Gierull e Sikaneta (2002). Além disso, tais estimativas podem fornecer um número de *looks* adequado a uma janela de dados selecionada.

No que segue, a próxima seção fornece detalhes sobre o modelo em discussão ($G\Gamma$).

2.1 Modelo $G\Gamma$

Nesta dissertação, consideramos um modelo mais geral que o apresentado anteriormente com fdp (2.2) para descrever intensidades I de $X(i, j)$. Sabe-se que a abordagem da *modelagem multiplicativa* discutida por Gao (2010) tem sido frequentemente utilizada devido à sua interpretabilidade física. No entanto, tal abordagem produz geralmente modelos estatísticos com fdps que possuem formas analíticas intratáveis (por exemplo, dependentes de funções especiais tal como a *função modificada de Bessel do terceiro tipo* da distribuição \mathcal{K} estudada por Blacknell (1994)). O uso da função especial impõe dificuldade à realização de inferência e proposta de métodos de teoria assintótica, objetivos

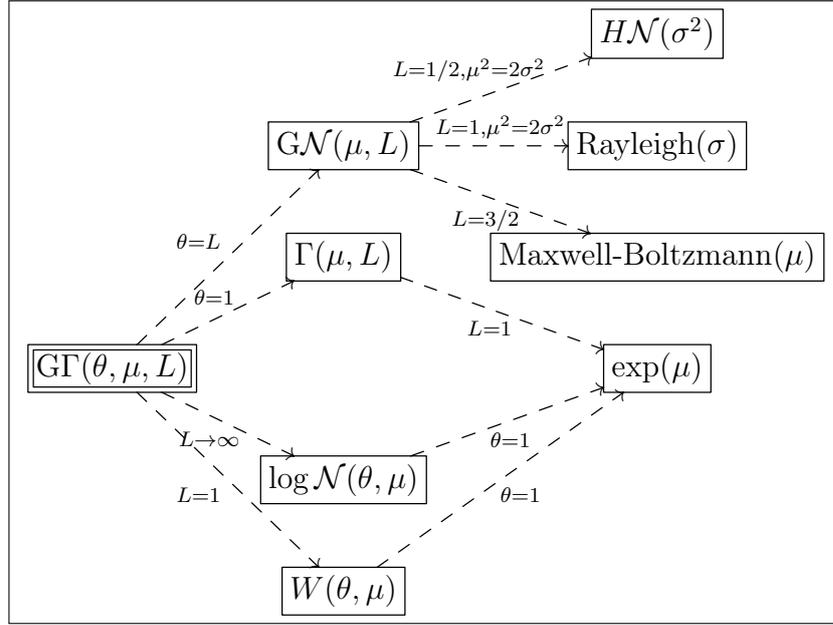


Figura 2.1: Diagrama ilustrando casos particulares do modelo $G\Gamma$ para dados de intensidade SAR.

secundários desta dissertação. Como alternativa, considerando a *abordagem de distribuição empírica* discutida por Gao (2010), sugere-se uma reparametrização da distribuição $G\Gamma$ proposta por Stacy (1962) tendo, sob modificação trivial, fdp dada por

$$f_Y(y; \theta, \mu, L) = \frac{\theta}{\mu \Gamma(L)} \left(\frac{y}{\mu}\right)^{\theta L - 1} \exp\left[-\left(\frac{y}{\mu}\right)^\theta\right] \mathbb{I}_{(0, \infty)}(y). \quad (2.3)$$

em que $\theta \in \mathbb{R}_+^*$. Esta situação é denotada por $Y \sim G\Gamma(\theta, \mu, L)$.

Para modelagem de dados de intensidade, consideramos a transformação $Z = L^{-1/\theta} Y$ como um modelo adequado. Assim, a variável aleatória Z tem fdp $f_Z(z; \bullet) = L^{1/\theta} \cdot f_Y(z \cdot L^{1/\theta}; \bullet)$ ou, equivalentemente,

$$f_Z(z; L, \mu, \theta) = \frac{|\theta| L^L}{\mu \Gamma(L)} \left(\frac{z}{\mu}\right)^{\theta L - 1} \exp\left[-L \left(\frac{z}{\mu}\right)^\theta\right] \mathbb{I}_{(0, \infty)}(z), \quad (2.4)$$

em que $\theta \in \mathbb{R}^*$. Usando a mesma notação de (2.3), tem-se que $Z \sim G\Gamma(\theta, \mu L^{-1/\theta}, L)$. A partir do modelo $G\Gamma$, várias distribuições importantes surgem como casos particulares, como ilustrado no diagrama apresentado na Figura 2.1. É evidente alguns casos particulares—tais como as distribuições normal generalizada (NG), log-normal (log N), half normal (HN), gama (Γ), Weibull (W), Rayleigh, Maxwell-Boltzmann e exponencial (exp)—sob restrições específicas e sob $\theta > 0$. Do mesmo modo, assumindo $\theta < 0$, são obtidas versões recíprocas desses casos particulares. Li *et al.* (2011) indicaram esta distribuição como uma modelagem empírica alternativa à modelagem multiplicativa para intensidades SAR. Nesse caso, os parâmetros θ , L , e μ foram considerados como os parâmetros de potência, forma e escala, respectivamente. Contudo, é possível notar que a

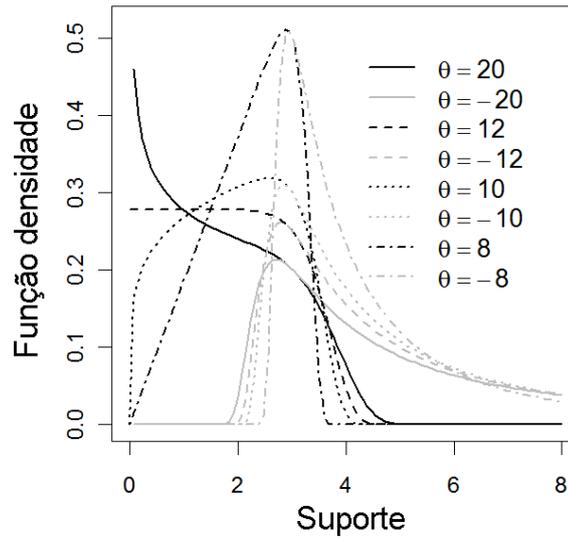


Figura 2.2: Curvas de densidade do modelo $G\Gamma$ para $[\theta, L, \mu] = [\bullet, 0.1, 3]$.

Equação (2.4) generaliza a Equação (2.2). Portanto, a Equação (2.4) também pode ser considerada como uma modelagem estendida para o ruído *speckle*, em que L e μ representam parâmetros do NEL e intensidade média, respectivamente; enquanto, θ pode ser entendido como um parâmetro de correção do número de looks para diferentes texturas de acordo com os experimentos que serão discutidos no Capítulo 4. Para ilustrar a flexibilidade da distribuição $G\Gamma$, várias curvas são apresentadas na Figura 2.2. Nota-se os seguintes comportamentos quando se especifica $[L, \mu] = [0.1, 3]$:

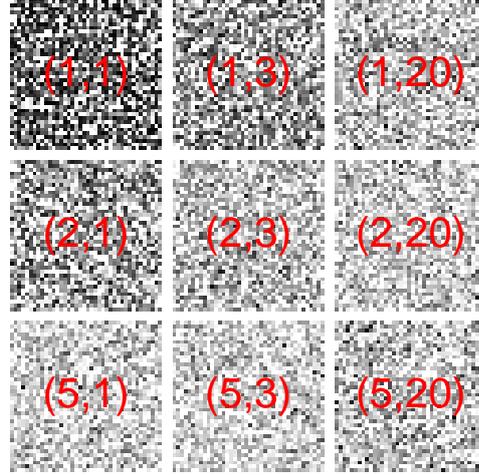
- Casos em que $\theta < 0$ são representados por fdps assimétricas à direita;
- O caso $\theta = 8$ produz uma curva assimétrica para a esquerda;
- Uma fdp degrau ocorre em $\theta = 12$ e
- Casos de bimodalidade leve podem ser caracterizados por $\theta = 10$ e 20 .

A Figura 2.3 apresenta uma imagem sintética gerada a partir da lei $G\Gamma$. Similarmente ao uso da distribuição Γ como modelo para o ruído *speckle*, esta figura ilustra que o aumento do parâmetro L melhora a relação sinal-ruído quando a lei $G\Gamma$ é considerada. Além disso, o aumento do parâmetro θ tende a descrever cenários de maior variabilidade.

A seguir, as Seções 2.2-2.3 discutem os métodos inferenciais utilizados nesta dissertação para o modelo $G\Gamma$ e apresentam algumas contribuições teóricas.

2.2 Inferência baseada em verossimilhança para o modelo $G\Gamma$

Seja Z uma variável aleatória tal que $Z \sim G\Gamma(\theta, \mu L^{-1/\theta}, L)$ e $\boldsymbol{\delta} = (\theta, L, \mu)^\top$ seu vetor de parâmetros, em que $(\cdot)^\top$ é o operador transposição. Então, a função de log-


 Figura 2.3: Imagem sintética para $L \in \{1, 3, 20\}$ e $\theta \in \{1, 2, 5\}$

verossimilhança associada a n observações, z_1, z_2, \dots, z_n , é

$$\begin{aligned} \ell_n(\boldsymbol{\delta}) \equiv \ell(\boldsymbol{\delta}; z_1, z_2, \dots, z_n) = & n \{ \log |\theta| + L \log L - L\theta \log \mu - \log \Gamma(L) \} \\ & + (L\theta - 1) \sum_{i=1}^n \log z_i - \frac{k}{\mu^\theta} \sum_{i=1}^n z_i^\theta, \end{aligned} \quad (2.5)$$

sendo o EMV de $\boldsymbol{\delta}$ determinado pela maximização de $\ell_n(\boldsymbol{\delta})$. Baseado na Equação (2.5), a função escore,

$$\mathbf{U}_{\boldsymbol{\delta}} = (U_\theta, U_L, U_\mu)^\top = \left(\frac{d \ell_n(\boldsymbol{\delta})}{d\theta}, \frac{d \ell_n(\boldsymbol{\delta})}{dL}, \frac{d \ell_n(\boldsymbol{\delta})}{d\mu} \right)^\top,$$

pode ser derivada da seguinte forma. Após algumas manipulações algébricas, os elementos da função escore $G\Gamma$ são dados por

$$U_\theta = \frac{n}{\theta} + L \sum_{i=1}^n \log z_i - n L \log \mu + \frac{L \log \mu}{\mu^\theta} \sum_{i=1}^n z_i^\theta - \frac{L}{\mu^\theta} \sum_{i=1}^n z_i^\theta \log z_i, \quad (2.6)$$

$$U_L = n \{ \log L + 1 - \theta \log \mu - \psi^{(0)}(L) \} + \theta \sum_{i=1}^n \log z_i - \frac{1}{\mu^\theta} \sum_{i=1}^n z_i^\theta, \quad (2.7)$$

e

$$U_\mu = -\frac{n L \theta}{\mu} + \frac{L \theta}{\mu^{\theta+1}} \sum_{i=1}^n z_i^\theta, \quad (2.8)$$

em que $\psi^{(k)}(x) = d^{k+1} \log \Gamma(x) / dx^{k+1}$, para $k = 0, 1, 2, \dots$, é a função poligama. De $U_\mu = 0$ e $U_\theta = 0$, tem-se que os EMVs para μ , $\hat{\mu}$, e L , \hat{L} , podem ser determinados em termos do

EMV para $\theta, \hat{\theta}$, pelas seguintes expressões:

$$\hat{\mu}(\theta) = \left(n^{-1} \sum_{i=1}^n z_i^\theta \right)^{1/\theta} \quad (2.9)$$

e

$$\hat{L}(\theta) = \frac{1}{\theta} \left[\frac{\sum_{i=1}^n z_i^\theta \log z_i}{\sum_{i=1}^n z_i^\theta} - n^{-1} \sum_{i=1}^n \log z_i \right]^{-1}. \quad (2.10)$$

Assim, o sistema de equações resultante de $\mathbf{U}_\theta = \mathbf{0}$ pode ser reduzido em encontrar uma solução em uma única equação não-linear dada por:

$$\log L(\hat{\theta}) - \log \left(n^{-1} \sum_{i=1}^n z_i^{\hat{\theta}} \right) - \psi^{(0)}(L(\hat{\theta})) + \hat{\theta} n^{-1} \sum_{i=1}^n \log z_i = 0. \quad (2.11)$$

Na prática, a quantidade $L(\hat{\theta})$ na expressão acima pode ser substituída pelo número de looks L (que é conhecido no processo de formação de imagem SAR), resultando em

$$\underbrace{\log L + 1 - \psi^{(0)}(L) + n^{-1} \sum_{i=1}^n \left\{ \hat{\theta} \log \frac{z_i}{\mu} - \left(\frac{z_i}{\mu} \right)^{\hat{\theta}} \right\}}_{\xi(L, \theta, \mu, \{z_1, \dots, z_n\})} = 0.$$

Note que $\xi(\cdot)|_{\theta=0} = \log L - \psi^{(0)}(L) \geq 0$ (a identidade ocorre para $L \rightarrow \infty$). Adicionalmente, com exceção do caso em que $z_i = \mu$ para $i = 1, 2, \dots, n$, existe $\theta_0 \in \mathbb{R}^*$ tal que $\xi(L, \theta, \mu, \{z_1, z_2, \dots, z_n\}) < 0$ para $\theta \geq \theta_0$ ou $\theta \leq \theta_0$, i.e., existe pelo menos uma raiz em $(0, \theta_0]$ ou $[\theta_0, 0)$, respectivamente. Portanto, pode-se definir a estimação pelo Algoritmo 1.

Algorithm 1 EMVs para os parâmetros do modelo GF em aplicações SAR e não-SAR.

- 1: **if** uma amostra observada $\{z_1, \dots, z_n\}$ é retirada de uma imagem SAR **then**
 - 2: calcular $\hat{\theta}$, sendo este uma solução da Equação (2.11), substituindo $L(\hat{\theta})$ por L ;
 - 3: calcular \hat{L} e $\hat{\mu}$ utilizando a solução de $\hat{\theta}$ em (2.9) e (2.10), respectivamente.
 - 4: **else**
 - 5: calcular $\hat{\theta}$, sendo este uma solução da Equação (2.11);
 - 6: calcular \hat{L} e $\hat{\mu}$ utilizando a solução de $\hat{\theta}$ em (2.9) e (2.10), respectivamente.
-

A partir das funções escore (2.6)–(2.8) acima, o seguinte resultado pode ser derivado por meio de algumas manipulações algébricas:

Teorema 1. *Seja $Z \sim GF(\theta, \mu L^{-1/\theta}, L)$, então os seguintes resultados são obtidos:*

- a) $E(Z^\theta) = \mu^\theta$
- b) $E(\log Z) = \frac{\{\theta \log \mu + \psi^{(0)}(L) - \log L\}}{\theta}$ e

$$c) E(Z^\theta \log Z) = \mu^\theta L^{-1} [\theta^{-1} + L E \log Z].$$

Este teorema é a primeira contribuição teórica desta dissertação. Para realizar inferência em dados $G\Gamma$ distribuídos, derivamos a matriz de informação de Fisher que é definida por $\mathcal{K}(\boldsymbol{\delta}) = E[d\ell_n(\boldsymbol{\delta})/d\boldsymbol{\delta} d\ell_n(\boldsymbol{\delta})/d\boldsymbol{\delta}^\top]$ e, satisfeitos os pressupostos do Lema 7.3.11 (pp. 338 Casella e Berger 2002), pode também ser expressa por

$$\mathcal{K}(\boldsymbol{\delta}) = -E(d^2 \ell_n(\boldsymbol{\delta})/d\boldsymbol{\delta} d\boldsymbol{\delta}^\top) = -E \begin{pmatrix} \frac{d^2 \ell_n(\boldsymbol{\delta})}{d\theta^2} & \frac{d^2 \ell_n(\boldsymbol{\delta})}{d\theta dL} & \frac{d^2 \ell_n(\boldsymbol{\delta})}{d\theta d\mu} \\ \bullet & \frac{d^2 \ell_n(\boldsymbol{\delta})}{dL^2} & \frac{d^2 \ell_n(\boldsymbol{\delta})}{dL d\mu} \\ \bullet & \bullet & \frac{d^2 \ell_n(\boldsymbol{\delta})}{d\mu^2} \end{pmatrix} = \begin{pmatrix} \kappa_{\theta,\theta} & \kappa_{\theta,L} & \kappa_{\theta,\mu} \\ \bullet & \kappa_{L,L} & \kappa_{L,\mu} \\ \bullet & \bullet & \kappa_{\mu,\mu} \end{pmatrix}.$$

Esta última identidade possui maior tratabilidade analítica. Em seguida, derivamos os elementos dessa importante matriz: Sejam Z_1, Z_2, \dots, Z_n uma amostra aleatória de $Z \sim G\Gamma(\theta, \mu L^{-1/\theta}, L)$,

$$\kappa_{\theta,\theta} = \frac{n}{\theta^2} + nL \log^2 \mu - \frac{2L \log \mu}{\mu^\theta} \sum_{i=1}^n E(Z_i^\theta \log Z_i) + \frac{L}{\mu^\theta} \sum_{i=1}^n E(Z_i^\theta \log^2 Z_i),$$

$$\kappa_{\theta,L} = \frac{1}{\mu^\theta} \sum_{i=1}^n E(Z_i^\theta \log Z_i) - \sum_{i=1}^n E(\log Z_i),$$

$$\kappa_{\theta,\mu} = \frac{n\theta L \log \mu}{\mu} - \frac{L\theta}{\mu^{\theta+1}} \sum_{i=1}^n E(Z_i^\theta \log Z_i),$$

$$\kappa_{L,L} = n\psi^{(1)}(L) - \frac{n}{L}, \quad \kappa_{L,\mu} = 0, \quad e \quad \kappa_{\mu,\mu} = \frac{nL\theta}{\mu^2},$$

em que todos os momentos são derivados no Teorema 1. Sob a satisfação das condições de regularidade (COX; HINKLEY, 1979), podemos aproximar a distribuição de $\sqrt{n}(\widehat{\boldsymbol{\delta}} - \boldsymbol{\delta})$ pela normal multivariada $N_3(0, \mathcal{K}(\boldsymbol{\delta})^{-1})$ e, portanto, realizar inferência assintótica. Por exemplo, verificar se θ em (2.4) é estatisticamente diferente de um (i.e. para testar a hipótese nula $H_0 : \theta = 1$ ou, equivalentemente, $H_0 : F_\Gamma = F_{G\Gamma}$, em que $F_{\mathcal{D}}$ representa a função de distribuição acumulada da distribuição \mathcal{D}) é um bom critério de decisão para comparar os modelos Γ e $G\Gamma$. Neste caso, a estatística da razão entre verossimilhanças (RV) é dada por

$$S_{RV} = 2 \{\ell(\widehat{\boldsymbol{\delta}}) - \ell(\widetilde{\boldsymbol{\delta}})\},$$

em que $\widehat{\boldsymbol{\delta}}$ e $\widetilde{\boldsymbol{\delta}}$ são os vetores dos EMVs irrestrito (sob H_1) e restrito (sob H_0) para $\boldsymbol{\delta}$, respectivamente. Sob a hipótese nula, a distribuição limite de S_{RV} é distribuição χ_1^2 . Se a estatística de teste exceder o quantil superior $100(1 - \alpha)\%$ da distribuição χ_1^2 , então rejeitamos a hipótese nula.

Na Seção 2.3 a seguir, aborda-se mais uma contribuição desta dissertação: um novo estimador para o NEL definido a partir da combinação entre autovalores da decompo-

sição espectral da matriz de coerência e das estimativas obtidas para NEL a partir da distribuição $G\Gamma$ em cada canal de polarização.

2.3 Refinamentos para inferência na distribuição $G\Gamma$: Estimador semiparamétrico para o NEL

As imagens PolSAR são tais que cada entrada é associada a quatro canais de polarização complexos, diga-se HH (horizontal-horizontal), HV (horizontal-vertical), VH (vertical-horizontal) e VV (vertical-vertical). Na prática, pelo Teorema da reciprocidade (ULABY; ELACHI, 1990), tem-se que $HV \approx VH$. Assim, a partir da perspectiva de dados *single-look*, cada pixel pode ser entendido como um resultado de um vetor aleatório, $\mathbf{y} = (S_{HH}, S_{HV}, S_{VV})^\top \in \mathbb{C}^3$. O vetor \mathbf{y} costuma ser modelado como $\mathbf{y} \sim N_3^{\mathbb{C}}(\mathbf{0}, \Sigma)$, em que $\Sigma = \text{Cov}(\mathbf{y})$ apresenta natureza hermitiana positiva definida. Embora a presença do ruído *speckle* seja inerente à captura de imagens SAR polarimétricas, a última distribuição não inclui um parâmetro associado ao controle de tal efeito.

A fim de superar esta dificuldade, usa-se o *processo multilook* definido da seguinte forma: Sejam $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_L$ uma amostra aleatória $\mathbf{Y}_k = (Y_k^{HH}, Y_k^{HV}, Y_k^{VV})^\top$ de $\mathbf{y} \sim N_3^{\mathbb{C}}(\mathbf{0}, \Sigma)$, o *processo multilook* é uma transformação nos dados *single-look* tal que a imagem resultante tenha seus *pixels* associados a uma matriz do tipo:

$$\mathbf{Z} = \frac{1}{L} \sum_{k=1}^L \mathbf{y}_k \mathbf{y}_k^* = \frac{1}{L} \sum_{k=1}^L \begin{bmatrix} \|Y_k^{HH}\|^2 & \langle Y_k^{HH}, Y_k^{HV} \rangle & \langle Y_k^{HH}, Y_k^{VV} \rangle \\ \langle Y_k^{HV}, Y_k^{HH} \rangle & \|Y_k^{HV}\|^2 & \langle Y_k^{HV}, Y_k^{VV} \rangle \\ \langle Y_k^{VV}, Y_k^{HH} \rangle & \langle Y_k^{VV}, Y_k^{HV} \rangle & \|Y_k^{VV}\|^2 \end{bmatrix},$$

em que $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* \mathbf{y}$ é o produto interno, $\|\mathbf{y}\| = \sqrt{\mathbf{y}^* \mathbf{y}}$ é a norma de qualquer argumento complexo, Y_k^M representa o canal M no k -ésimo *look* e $*$ indica o operador transposição conjugado. De acordo com Goodman (1963), tem-se que \mathbf{Z} segue a distribuição Wishart complexa escalonada (WCE) com fdp dada por

$$f_{\mathbf{Z}}(\mathbf{Z}; L, \Sigma) = \frac{L^p}{|\Sigma|^L \Gamma_p(L)} |\mathbf{Z}|^{L-p} \exp \left\{ -L \text{tr}(\Sigma^{-1} \mathbf{Z}) \right\}, \quad (2.12)$$

em que $\Gamma_p(L) = \prod_{k=0}^{p-1} \Gamma(L-k)$ é a função gama multivariada, $|\cdot|$ é o operador determinante, $\text{tr}(\cdot)$ é o operador traço, $\Sigma = E\{\mathbf{Z}\}/L$ e p é a quantidade de canais de polarização. Esta lei é denotada como $\mathbf{Z} \sim \mathcal{W}_3^{\mathbb{C}}(L, \Sigma)$.

Com base em (2.12) e assumindo que os dados PolSAR provém da matriz de coerência $\mathbf{C} = L^{-1} \mathbf{Z}$, Anfinsen *et al.* (2009) propuseram um estimador para o NEL. Sejam $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n$ uma amostra aleatória de $\mathbf{Z} \sim \mathcal{W}_3^{\mathbb{C}}(L, \Sigma)$. A partir de

$$\frac{\partial \sum_{k=1}^n \log f(\mathbf{Z}_k; L, \Sigma)}{\partial L} = 0,$$

o EMV para L pode ser dado pela solução da seguinte equação não linear:

$$g(\widehat{L}_e^{(\text{ML})}) = \langle \log |C| \rangle - \log |\langle C \rangle| - \sum_{i=0}^{d-1} \psi^{(0)}(\widehat{L}_e^{(\text{ML})} - i) + d \log \widehat{L}_e^{(\text{ML})} = 0, \quad (2.13)$$

em que $\langle \log |C| \rangle = n^{-1} \sum_{k=1}^n \log |L^{-1} Z_k|$ e $\langle C \rangle = n^{-1} \sum_{k=1}^n L^{-1} Z_k$.

Nesta dissertação, propomos um novo estimador para o NEL comparável aquele proposto por Anfinsen *et al.* (2009), porém de implementação trivial, baixo custo computacional e menor viés, conforme será abordado no capítulo 4. Ele é definido a partir da combinação entre autovalores da decomposição espectral do retorno de um pixel e EMVs para o NEL a partir das intensidades dos canais de polarização HH, HV e VV. Os novos estimadores são definidos como

$$\widehat{L}(i, j) = C^{-1}(i, j) \sum_{k=1}^3 \lambda_k(i, j) \widehat{L}_{\text{GT}}^{(k)}(i, j). \quad (2.14)$$

em que $C(i, j) = \text{tr}[\Sigma(i, j)] = \sum_{k=1}^3 \lambda_k(i, j)$, $\lambda_k(i, j)$ é o autovalor dominante definido pelo k -ésimo canal da entrada (i, j) e $\widehat{L}_{\text{GT}}^{(k)}(i, j)$ é o EMV de L baseado em uma amostra ao redor do pixel (i, j) associada ao k -ésimo canal. No que segue, ilustra-se a definição de autovalor dominante.

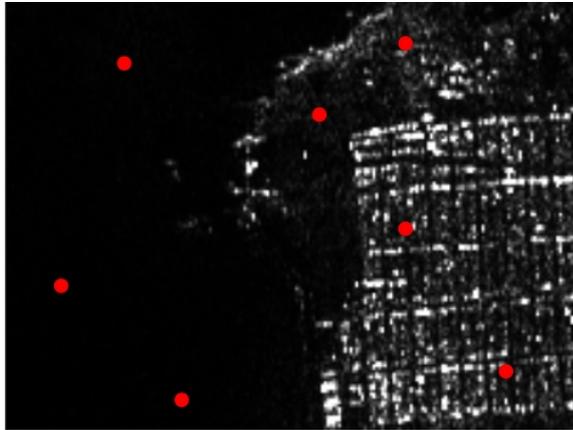


Figura 2.4: Recorte da imagem AIRSAR de São Francisco com *pixels* destacados.

A Figura 2.4 apresenta um recorte da imagem AIRSAR da Baía de São Francisco, obtida assumindo um número de *looks* igual a quatro.

Realizamos um primeiro experimento para analisar empiricamente se cada canal apresenta um autovalor diferente quando se considera a decomposição espectral (DE) das matrizes de coerência devido aos pixels destacados na Figura 2.4. A Tabela 2.1 apresenta informações sobre a DE dos sete pixels na Figura 2.1 no seguinte sentido. Sabe-se que todas as matrizes quadradas complexas podem ser decompostas pela decomposição de Gauss-Jordan (que generaliza a DE) e, portanto, a coerência pode ser representada como

Tabela 2.1: Tabela *Pixel-M* baseada em recorte da imagem AIRSAR de São Francisco.

Pixel	Canal	C_1^M	C_2^M	C_3^M	$\lambda_1 > \lambda_2 > \lambda_3$
1	HH	0.1972	0.0230	0.7798	0.0001
	HV	0.2556	0.7311	0.0132	0.0027
	VV	0.5471	0.2459	0.2069	0.0350
2	HH	0.1507	0.0021	0.8470	0.000097
	HV	0.1440	0.8205	0.0353	0.000930
	VV	0.7051	0.1772	0.1175	0.016000
3	HH	0.3397	0.0095	0.6506	0.0002
	HV	0.6402	0.03013	0.3296	0.0089
	VV	0.0200	0.9602	0.0196	0.0030
4	HH	0.7356	0.2453	0.0190	0.0434
	HV	0.0952	0.2026	0.7020	0.0010
	VV	0.1691	0.5520	0.2788	0.0144
5	HH	0.4264	0.4875	0.0859	0.1137
	HV	0.5520	0.3895	0.0584	0.2423
	VV	0.0214	0.1229	0.8556	0.0081
6	HH	0.1940	0.0689	0.7369	0.0153
	HV	0.7312	0.0369	0.2317	0.4235
	VV	0.0746	0.8940	0.0313	0.1456
7	HH	0.2145	0.0081	0.7772	0.0044
	HV	0.7254	0.0814	0.1931	0.1177
	VV	0.0599	0.9104	0.0295	0.0401

$$\mathbf{Z} = \mathbf{T} \mathbf{\Lambda} \mathbf{T}^{-1}$$

em que \mathbf{T} e $\mathbf{\Lambda}$ são as matrizes de autovetores e autovalores (diagonal) de \mathbf{Z} , respectivamente. Assim, a entrada- (i, i) de \mathbf{Z} (ou o i -ésimo elemento da diagonal principal de \mathbf{Z}) pode ser escrita como

$$Z(i, i) = \sum_{i=1}^3 \overbrace{t_{1i} t^{i1}}^{c_i} \lambda_i, \quad (2.15)$$

em que t_{ij} e t^{ij} são as entradas (i, j) de \mathbf{T} e \mathbf{T}^{-1} , respectivamente. Definimos como *autovalores dominantes* de Z o autovalor associado com o maior valor c_i em (2.15). Na Tabela 2.1, pode-se notar que há apenas um autovalor dominante para cada canal.

Apresentadas as propostas sobre estimação por MV para os parâmetros da distribuição $\Gamma\Gamma$, incluindo-se o estimador semi-paramétrico, sobretudo a interpretabilidade física deste modelo quando aplicado a dados de intensidade SAR, no que segue, propõem-se estimadores corrigidos para os parâmetros $\Gamma\Gamma$ por meio da proposta de Cox-Snell.

2.4 Estimador corrigido segundo Cox-Snell para os parâmetros do modelo $\Gamma\Gamma$

Os EMV são frequentemente viesados para os verdadeiros valores dos parâmetros em modelos não lineares para pequenos tamanhos amostrais ou quando a informação de Fisher é reduzida, tendo viés $O(n^{-1})$. Este não é considerado em situações práticas, pois

seu valor é desprezível em relação ao erro padrão, que está em $O(n^{-1/2})$ (CORDEIRO, 2011). Entretanto, em amostras onde o tamanho não é suficientemente grande o viés pode ser substancial, apesar de não representar um problema sério em amostras grandes.

Devido a relevância do EMV, técnicas foram desenvolvidas para corrigir o viés destes estimadores em pequenas amostras. Em modelos uniparamétricos, Bartlett (1953) deduziu uma fórmula para o viés que está em $O(n^{-1})$ da EMV no caso em que as observações são independentes e identicamente distribuídas. Nos modelos multiparamétricos, os vieses de ordem n^{-1} das EMV, supondo uma amostra independente mas não necessariamente identicamente distribuída foram deduzidos em generalidade por Cox e Snell (1968). Para correção do viés de segunda ordem do EMV, Cox e Snell (1968) também propuseram uma fórmula geral.

Vários trabalhos foram realizados com objetivo de fornecer expressões para os vieses de segunda ordem dos EMV, entre eles, Lopes (2007) propôs expressões para os parâmetros do modelo de regressão normal assimétrico, utilizando-as para obter estimadores corrigidos, enquanto Dourado (2004) apresentou técnicas para remover o viés de segunda ordem para EMVs na família exponencial biparamétrica através de três procedimentos.

Consideramos inicialmente uma breve discussão sobre correção de viés, em que as derivadas da log-verossimilhança desempenham um papel importante na estimação MV. Então, assumamos (COX; SNELL, 1968; CORDEIRO, 2011):

$$U_{\theta_i} = \frac{\partial}{\partial \theta_i} \ell(\boldsymbol{\theta}), \quad U_{\theta_i \theta_j} = \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell(\boldsymbol{\theta}) \quad \text{e} \quad U_{\theta_i \theta_j \theta_k} = \frac{\partial^3}{\partial \theta_i \partial \theta_j \partial \theta_k} \ell(\boldsymbol{\theta}),$$

para $i, j, k = 1, 2, \dots, p$, em que p é a dimensão do vetor de parâmetros. Com base na definição acima os cumulantes são dados por:

$$\kappa_{\theta_i \theta_j} = E(U_{\theta_i \theta_j}), \quad \kappa_{\theta_i, \theta_j} = E(U_{\theta_i} U_{\theta_j}), \quad \kappa_{\theta_i \theta_j \theta_k} = E(U_{\theta_i \theta_j \theta_k}) \quad \text{e} \quad \kappa_{\theta_i, \theta_j \theta_k} = E(U_{\theta_i} U_{\theta_j \theta_k}).$$

Neste caso, a entrada (i, j) da matriz de informação de Fisher associada a $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_m]^\top$ é $\kappa_{\theta_i, \theta_j}$. Adicionalmente, denotamos os elementos da inversa da matriz de informação de Fisher como $\kappa^{\theta_i, \theta_j}$. Defina ainda,

$$\kappa_{\theta_i \theta_j}^{(\theta_k)} = \frac{\partial}{\partial \theta_k} \kappa_{\theta_i \theta_j},$$

para $i, j, k = 1, 2, \dots, p$. Agora, estamos em posição de definir a expressão de Cox-Snell:

$$B(\widehat{\boldsymbol{\theta}}_r) = \sum_{s, t, u} \kappa^{\theta_r, \theta_s} \kappa^{\theta_t, \theta_u} \left(\kappa_{\theta_s \theta_t}^{(\theta_u)} - \frac{1}{2} \kappa_{\theta_s \theta_t \theta_u} \right), \quad (2.16)$$

para $r = 1, 2, \dots, p$. Portanto, o EMV corrigido $\widetilde{\boldsymbol{\theta}}_r$ é definido por

$$\widetilde{\boldsymbol{\theta}}_r = \widehat{\boldsymbol{\theta}}_r - \widehat{B}(\widehat{\boldsymbol{\theta}}_r), \quad (2.17)$$

em que $\widehat{B}(\widehat{\theta}_r)$ representa o viés $B(\widehat{\theta}_r)$ avaliado em $\widehat{\theta}_r$. Adicionalmente, pode-se mostrar que (i) $E[\widehat{B}(\widehat{\theta}_r)] = O(N^{-2})$, (ii) $E(\widehat{\theta}_r) = \theta_r + O(N^{-1})$, e (iii) $E(\widetilde{\theta}_r) = \theta_r + O(N^{-2})$ (COX; SNELL, 1968). Portanto, o viés de $\widetilde{\theta}_r$ tem ordem de N^{-2} com $B(\widehat{\theta}_r) = O(N^{-1})$. Assim, espera-se que $\widetilde{\theta}_r$ possua melhores propriedades assintóticas quando comparado com $\widehat{\theta}_r$. No que segue, apresentam-se expressões para $B(\theta_i)$ para os modelos $G\Gamma$ escalonado e não-escalonado.

2.4.1 Viés para modelo $G\Gamma$ escalonado

Por questões de simplicidade analítica, assumimos $\theta > 0$, em (2.4):

$$f_Z(z; L, \mu, \theta) = \theta \frac{L^L}{\mu \Gamma(L)} \left(\frac{z}{\mu}\right)^{\theta L - 1} \exp\left[-L\left(\frac{z}{\mu}\right)^\theta\right] \mathbb{I}_{(0, \infty)}(z). \quad (2.18)$$

Seja uma amostra aleatória de tamanho N de $Z \sim G\Gamma(\theta, \mu L^{-1/\theta}, \theta)$, $\mathbf{z} = (z_1, z_2, \dots, z_N)^\top$, a função de log-verossimilhança é

$$\ell_n(\boldsymbol{\delta}) = N[\log \theta + L \log L - \log \mu - \log \Gamma(L)] + (\theta L - 1) \sum_{i=1}^N \log \frac{z_i}{\mu} - L \sum_{i=1}^N \left(\frac{z_i}{\mu}\right)^\theta. \quad (2.19)$$

em que $\boldsymbol{\delta} = (L, \theta)^\top$. Aqui, o parâmetro μ será entendido como uma quantidade conhecida dada pela média das observações de intensidade da amostra. Com base na Equação (2.19), a função escore,

$$\mathbf{U}_\delta = (U_\theta, U_L)^\top = \left(\frac{d \ell_n(\boldsymbol{\delta})}{d\theta}, \frac{d \ell_n(\boldsymbol{\delta})}{dL} \right)^\top,$$

pode ser derivada como se segue. Os estimadores para L e θ , diga-se \widehat{L} e $\widehat{\theta}$, respectivamente, são definidos como a solução do sistema de equações não lineares $\mathbf{U}_\delta = \mathbf{0}$, ou equivalentemente,

$$\begin{cases} N \widehat{\theta}^{-1} + \widehat{L} \sum_{i=1}^N \log \frac{z_i}{\mu} - \widehat{L} \sum_{i=1}^N \left(\frac{z_i}{\mu}\right)^{\widehat{\theta}} \log \frac{z_i}{\mu} = 0, \\ N \log \widehat{L} + N - N \psi^{(0)}(\widehat{L}) + \widehat{\theta} \sum_{i=1}^N \log \frac{z_i}{\mu} - \sum_{i=1}^N \left(\frac{z_i}{\mu}\right)^{\widehat{\theta}} = 0. \end{cases} \quad (2.20)$$

Observe que uma expressão em forma fechada para estimativa MV do NEL condicionada ao parâmetro θ conhecido é dada por

$$\widehat{L}(\theta) = \frac{N}{\theta} \left[\sum_{i=1}^N \left(\frac{z_i}{\mu}\right)^\theta \log \left(\frac{z_i}{\mu}\right) - \sum_{i=1}^N \log \left(\frac{z_i}{\mu}\right) \right]^{-1}.$$

Assim, o método de estimação MV é simplificado por encontrar a solução para a equação não-linear

$$\log \widehat{L}(\widehat{\theta}) + 1 - \psi^0(\widehat{L}(\widehat{\theta})) + N^{-1} \sum_{i=1}^N \left\{ \widehat{\theta} \log \frac{z_i}{\mu} - \left(\frac{z_i}{\mu} \right)^{\widehat{\theta}} \right\} = 0$$

Na prática, $\widehat{L}(\widehat{\theta})$ na expressão acima pode ser substituído pelo número de *looks* L (que é conhecido no processo de formação de imagens SAR) resultando em

$$\underbrace{\log L + 1 - \psi^0(L) + N^{-1} \sum_{i=1}^N \left\{ \widehat{\theta} \log \frac{z_i}{\mu} - \left(\frac{z_i}{\mu} \right)^{\widehat{\theta}} \right\}}_{\xi(L, \theta, \mu, \{z_1, \dots, z_N\})} = 0$$

Para obter a expressão em forma fechada para a matriz de informação de Fisher \mathcal{M}_δ ,

$$\mathcal{M}_\delta = \mathbb{E} \left(\frac{d\ell(\boldsymbol{\delta})}{d\boldsymbol{\delta}} \frac{d\ell(\boldsymbol{\delta})}{d\boldsymbol{\delta}^\top} \right) = -\mathbb{E} \left(\frac{d^2 \ell_n(\boldsymbol{\delta})}{d\boldsymbol{\delta} d\boldsymbol{\delta}^\top} \right) = -\mathbb{E} \begin{pmatrix} \frac{d^2 \ell(\boldsymbol{\delta})}{d\theta^2} & \frac{d^2 \ell(\boldsymbol{\delta})}{d\theta dL} \\ \bullet & \frac{d^2 \ell(\boldsymbol{\delta})}{dL^2} \end{pmatrix} = \begin{pmatrix} \kappa_{\theta, \theta} & \kappa_{\theta, L} \\ \bullet & \kappa_{L, L} \end{pmatrix},$$

derivamos o seguinte teorema.

Teorema 2. *Seja $Z \sim GF(L, \mu L^{-1/\theta}, \theta)$, então os seguintes valores esperados são válidos:*

- a) $\mathbb{E} \left\{ \log \left(\frac{Z}{\mu} \right) \right\} = \theta^{-1} [\psi^{(0)}(L) - \log(L)];$
- b) $\mathbb{E} \left\{ \left(\frac{Z}{\mu} \right)^\theta \log \left(\frac{Z}{\mu} \right) \right\} = \theta^{-1} [-\log(L) + L^{-1} + \psi^{(0)}(L)];$
- c) $\mathbb{E} \left\{ \left(\frac{Z}{\mu} \right)^\theta \right\} = 1;$
- d)

$$\begin{aligned} \mathbb{E} \left\{ \left(\frac{Z}{\mu} \right)^\theta \log^2 \left(\frac{Z}{\mu} \right) \right\} &= \theta^{-2} \log^2 L - \frac{2 \log L}{\theta L} \left\{ \frac{1 + L \psi^{(0)}(L)}{\theta} \right\} \\ &\quad + L^{-1} \theta^{-2} \{ 2 \psi^{(0)}(L) + L \psi^{(0)}(L)^2 + L \psi^{(1)}(L) \}; \end{aligned}$$

e)

$$\begin{aligned} E \left\{ \left(\frac{Z}{\mu} \right)^\theta \log^3 \left(\frac{Z}{\mu} \right) \right\} &= -\theta^{-3} \log^3 L + 3L^{-1} \theta^{-2} \log^2 L \left\{ \frac{1 + L\psi^{(0)}(L)}{\theta} \right\} \\ &\quad - 3L^{-1} \theta^{-1} \log L \left\{ \frac{2\psi^{(0)}(L) + L\psi^{(0)}(L)^2 + L\psi^{(1)}(L)}{\theta^2} \right\} \\ &\quad + L^{-1} \left\{ \frac{3\psi^{(0)}(L)^2 + L\psi^{(0)}(L)^3 + 3L\psi^{(0)}(L)\psi^{(1)}(L)}{\theta^3} \right. \\ &\quad \left. + \frac{3\psi^{(1)}(L) + L\psi^{(2)}(L)}{\theta^3} \right\}. \end{aligned}$$

Finalmente, temos que

$$\frac{\mathcal{M}_\delta}{N} = \begin{pmatrix} \theta^{-2} + LA_{L,\theta} & (\theta L)^{-1} \\ \bullet & \psi^{(1)}(L) - L^{-1} \end{pmatrix},$$

em que $A_{L,\theta} = E\{(Z/\mu)^\theta \log^2(Z/\mu)\}$ é dado no Teorema 2. O desenvolvimento dos demais elementos de (2.16) é dado a seguir. Com base no Teorema 2, derivamos as seguintes quantidades:

(i) Elementos da inversa da matriz de informação de Fisher:

$$\kappa^{\theta,\theta} = N^{-1} D_{L,\theta}^{-1} [\psi^{(1)}(L) - L^{-1}], \quad \kappa^{L,\theta} = -N^{-1} [\theta L D_{L,\theta}]^{-1}$$

e

$$\kappa^{L,L} = N^{-1} D_{L,\theta}^{-1} [\theta^{-2} + LA_{L,\theta}],$$

em que

$$D_{L,\theta} = \theta^{-2} [\psi^{(1)}(L) - L^{-1}] - (\theta L)^{-2} + A_{L,\theta} [L\psi^{(1)}(L) - 1].$$

(ii) Cumulantes baseados nas terceiras derivadas da função de log-verossimilhança:

$$\kappa_{\theta\theta}^{(\theta)} = N \left[\frac{2}{\theta^3} - LA_{L,\theta}^{(\theta)} \right], \quad \kappa_{\theta\theta}^{(L)} = -N \left[A_{L,\theta} + LA_{L,\theta}^{(L)} \right],$$

$$\kappa_{L\theta}^{(\theta)} = \kappa_{\theta L}^{(\theta)} = \frac{N}{L\theta^2}, \quad \kappa_{L\theta}^{(L)} = \kappa_{\theta L}^{(L)} = \frac{N}{L^2\theta}, \quad \kappa_{LLL} = \kappa_{LL}^{(L)} = -N [L^{-2} + \psi^{(2)}(L)],$$

$$\kappa_{L\theta\theta} = \kappa_{\theta\theta L} = \kappa_{\theta L\theta} = -NA_{L,\theta}, \quad \kappa_{\theta\theta\theta} = \frac{2N}{\theta^3} - NL C_{L,\theta},$$

$$\kappa_{LL\theta} = \kappa_{L\theta L} = \kappa_{\theta LL} = \kappa_{LL}^{(\theta)} = 0,$$

em que

$$C_{L,\theta} = E \left\{ \left(\frac{Z}{\mu} \right)^\theta \log^3 \left(\frac{Z}{\mu} \right) \right\}, \quad A_{L,\theta}^{(\theta)} = -2\theta^{-1} A_{L,\theta}$$

e

$$\begin{aligned} A_{L,\theta}^{(L)} &= \frac{2 \log L}{\theta^2 L} [1 - \psi^{(0)}(L) - L\psi^{(1)}(L)] - 2 \left[\frac{1 - \log L}{(\theta L)^2} \right] \left[1 + L\psi^{(0)}(L) \right] \\ &\quad + \frac{1}{L\theta^2} \left\{ 2\psi^{(1)}(L) - 2L^{-1}\psi^{(0)}(L) + 2L\psi^{(0)}(L)\psi^{(1)}(L) + L\psi^{(2)}(L) \right\}. \end{aligned}$$

Assim, expressões fechadas para $B(\theta)$ e $B(L)$ são obtidas das quantidades acima:

$$\begin{aligned} B(\hat{\theta}) &= \frac{D_{L,\theta}^{-2}}{N} [\psi^{(1)}(L) - L^{-1}]^2 \left[\theta^{-3} + \frac{LC_{L,\theta}}{2} - LA_{L,\theta}^{(\theta)} \right] \\ &\quad + \frac{L^{-1}\theta^{-1}D_{L,\theta}^{-2}}{N} [\psi^{(1)}(L) - L^{-1}] \left(\frac{A_{L,\theta}}{2} - L^{-1}\theta^{-2} + LA_{L,\theta}^{(L)} \right) \\ &\quad + \frac{L^{-4}\theta^{-3}D_{L,\theta}^{-2}}{N} \end{aligned} \quad (2.21)$$

e

$$\begin{aligned} B(\hat{L}) &= - \frac{\theta^{-2} L^{-2} D_{L,\theta}^{-2}}{N} \left[L^{-1}\theta^{-2} + 2A_{L,\theta} + LA_{L,\theta}^{(L)} \right] \\ &\quad - \frac{D_{L,\theta}^{-2}}{N} \left[\frac{\theta^{-1}C_{L,\theta}}{2} - \theta^{-1}A_{L,\theta}^{(\theta)} - \frac{3\theta^{-1}A_{L,\theta}}{2} - \frac{L(A_{L,\theta})^2}{2} \right]. \end{aligned} \quad (2.22)$$

2.4.2 Viés para modelo $G\Gamma$ não escalonado

Analogamente ao desenvolvimento realizado para se obter a expressão em forma fechada para a matriz de informação de Fisher \mathcal{M}_δ sob a distribuição $G\Gamma$ escalonada, temos o seguinte teorema.

Teorema 3. *Seja $Y \sim G\Gamma(\theta, \mu, L)$, então os seguintes valores esperados são válidos:*

$$a) \ E \left\{ \log \left(\frac{Y}{\mu} \right) \right\} = \theta^{-1} \psi^{(0)}(L);$$

$$b) \ E \left\{ \left(\frac{Y}{\mu} \right)^\theta \log \left(\frac{Y}{\mu} \right) \right\} = \theta^{-1} [1 + L \psi^{(0)}(L)];$$

$$c) \ E \left\{ \left(\frac{Y}{\mu} \right)^\theta \right\} = L;$$

d)

$$E \left\{ \left(\frac{Y}{\mu} \right)^\theta \log^2 \left(\frac{Y}{\mu} \right) \right\} = \theta^{-2} [L \psi^{(0)}(L)^2 + 2 \psi^{(0)}(L) + L \psi^{(1)}(L)];$$

e)

$$E \left\{ \left(\frac{Y}{\mu} \right)^\theta \log^3 \left(\frac{Y}{\mu} \right) \right\} = \theta^{-3} [3 \psi^{(0)}(L)^2 + 3 \psi^{(1)}(L) + L \psi^{(0)}(L)^3] + \theta^{-3} [3L \psi^{(0)}(L) \psi^{(1)}(L) + L \psi^{(2)}(L)]$$

Finalmente, temos que

$$\frac{\mathcal{M}_\delta}{N} = \begin{pmatrix} \theta^{-2} + A_{L,\theta} & -\psi^{(0)}(L)\theta^{-1} \\ \bullet & \psi^{(1)}(L) \end{pmatrix},$$

em que $A_{L,\theta} = E\{(Y/\mu)^\theta \log^2(Y/\mu)\}$ é dado no Teorema 3. O desenvolvimento dos demais elementos de (2.16) é dado a seguir. Com base no Teorema 3, derivamos as seguintes quantidades:

(i) Elementos da inversa da matriz de informação de Fisher:

$$\kappa^{\theta,\theta} = N D_{L,\theta}^{-1} \psi^{(1)}(L) \quad \kappa^{L,\theta} = N \psi^{(0)}(L) [\theta D_{L,\theta}]^{-1}$$

e

$$\kappa^{L,L} = N D_{L,\theta}^{-1} [\theta^{-2} + A_{L,\theta}],$$

em que

$$D_{L,\theta} = \psi^{(1)}(L) [\theta^{-2} + A_{L,\theta}] - [\psi^{(0)}(L) \theta^{-1}]^{-2}.$$

(ii) Cumulantes baseados nas terceiras derivadas da função de log-verossimilhança:

$$\kappa_{\theta\theta}^{(\theta)} = N \left[\frac{2}{\theta^3} - A_{L,\theta}^{(\theta)} \right], \quad \kappa_{\theta\theta}^{(L)} = -N A_{L,\theta}^{(L)}, \quad \kappa_{L\theta}^{(\theta)} = \kappa_{\theta L}^{(\theta)} = -\frac{N \psi^{(0)}(L)}{\theta^2},$$

$$\kappa_{L\theta}^{(L)} = \kappa_{\theta L}^{(L)} = \frac{N \psi^{(1)}(L)}{\theta}, \quad \kappa_{LLL} = \kappa_{LL}^{(L)} = -N \psi^{(2)}(L),$$

$$\kappa_{LL\theta} = \kappa_{L\theta L} = \kappa_{\theta LL} = \kappa_{L\theta\theta} = \kappa_{\theta\theta L} = \kappa_{\theta L\theta} = \kappa_{LL}^{(\theta)} = 0,$$

$$\kappa_{\theta\theta\theta} = \frac{2N}{\theta^3} - N C_{L,\theta},$$

em que

$$C_{L,\theta} = \mathbb{E} \left\{ \left(\frac{Y}{\mu} \right)^\theta \log^3 \left(\frac{Y}{\mu} \right) \right\}, \quad A_{L,\theta}^{(\theta)} = -2 \theta^{-1} A_{L,\theta}$$

e

$$A_{L,\theta}^{(L)} = \theta^{-2} \left\{ 3 \psi^{(1)}(L) + \psi^{(0)}(L)^2 + 2L \psi^{(0)}(L) \psi^{(1)}(L) + L \psi^{(2)}(L) \right\}.$$

Finalmente, derivamos expressões fechadas para $B(\theta)$ e $B(L)$ dadas pelas equações:

$$\begin{aligned} B(\widehat{\theta}) = & D_{L,\theta}^{-2} N^3 \left\{ \psi^{(1)}(L)^2 \left[\theta^{-3} - A_{L,\theta}^{(\theta)} + \frac{1}{2} C_{L,\theta} \right] - \psi^{(1)}(L) \psi^{(0)}(L)^2 \theta^{-3} \right. \\ & \left. + \psi^{(1)}(L) \psi^{(0)}(L) \theta^{-1} \left[-A_{L,\theta}^{(L)} \right] \right\} + D_{L,\theta}^{-2} N^3 \left\{ \psi^{(1)}(L)^2 \theta^{-1} [\theta^{-2} + A_{L,\theta}] \right. \\ & \left. - \frac{1}{2} \psi^{(0)}(L) \psi^{(2)}(L) \theta^{-1} [\theta^{-2} + A_{L,\theta}] \right\} \end{aligned} \quad (2.23)$$

e

$$\begin{aligned} B(\widehat{L}) = & D_{L,\theta}^{-2} N^3 \left\{ -\frac{1}{2} \psi^{(2)}(L) [\theta^{-2} + A_{L,\theta}]^2 + \psi^{(0)}(L) \psi^{(1)}(L) \theta^{-2} [\theta^{-2} + A_{L,\theta}] \right. \\ & \left. - \psi^{(0)}(L)^3 \theta^{-4} \right\} + D_{L,\theta}^{-2} N^3 \left\{ -[\psi^{(0)}(L) \theta^{-1}]^2 A_{L,\theta}^{(L)} \right. \\ & \left. + \psi^{(0)}(L) \psi^{(1)}(L) \theta^{-1} \left[\theta^{-3} + \frac{C_{L,\theta}}{2} - A_{L,\theta}^{(\theta)} \right] \right\}. \end{aligned} \quad (2.24)$$

O próximo capítulo apresenta um tratamento estatístico para testes de hipóteses baseados em medidas de divergência e distância de Kullback-Leibler conforme a classe de divergências (h, ϕ) proposta por Salicrú *et al.* (1994).

Teoria da Informação Estatística: Divergência Estocástica para os
modelos Γ e $\mathcal{W}^{\mathbb{C}}$

A Teoria da Informação tem fornecido ferramentas fundamentais para aplicações em inferência estatística (BLATT; III, 2007), compressão de dados (DONOHO *et al.*, 1998) e processamento de imagem (MORIO *et al.*, 2009). A idéia inicial sobre esta teoria foi proposta por Hartley na década de 1930 (HARTLEY, 1928), ao introduzir uma medida logarítmica de informação para comunicação em estudos técnicos sobre a transmissão de dados em telégrafos. Posteriormente, Shannon (SHANNON, 1948) formalizou essa ideia, definindo os conceitos matemáticos de entropia e informação mútua no contexto de comunicações de dados. Em particular, a entropia pode ser entendida como um conceito basilar relacionado à noção de desordem em Mecânica Estatística (KULLBACK, 1968).

Finalmente, a primeira noção de entropia relativa (posteriormente cunhada como *divergência*) foi desenvolvida por Kullback e Leibler (KULLBACK; LEIBLER, 1951). A medida de Kullback-Leibler pode ser entendida como um critério de comparação entre duas distribuições de probabilidade. Em Estatística, surge como um logaritmo esperado da razão de verossimilhança: Sejam Z e Y duas variáveis aleatórias com fdps $f_Z(z; \boldsymbol{\theta}_z)$ e $f_Y(z; \boldsymbol{\theta}_y)$, em um suporte comum $\mathcal{X} \subset \mathbb{R}$. Então, a divergência de Y com respeito a Z é definida por

$$D_{\text{KL}}(Z \parallel Y) = \mathbb{E}_Z \left[\log \frac{f_Z(Z; \boldsymbol{\theta}_z)}{f_Y(Z; \boldsymbol{\theta}_y)} \right] = \int_{\mathcal{X}} f_Z(z; \boldsymbol{\theta}_z) \log \frac{f_Z(z; \boldsymbol{\theta}_z)}{f_Y(z; \boldsymbol{\theta}_y)} dz, \quad (3.1)$$

em que $\mathbb{E}_Z[h(Z)] = \int_{\mathcal{X}} h(z) f_Z(z; \boldsymbol{\theta}_z) dz$. É notório que Equação (3.1) é sempre não-negativa e assume o valor zero se, e somente se, $f_Z(z; \boldsymbol{\theta}_z) = f_Y(z; \boldsymbol{\theta}_y)$ para todo $z \in \mathcal{X}$. A divergência $D_{\text{KL}}(Z \parallel Y)$ é uma quantidade que expressa a ineficiência de assumir a distribuição de Y quando o modelo Z é verdadeiro (COVER; THOMAS, 1991). Embora

$D_{\text{KL}}(\cdot \parallel \cdot)$ seja chamada de ‘distância’ em alguns trabalhos, ela não é simétrica e, portanto, não pode ser considerada uma medida de distância. Nesta dissertação, considera-se uma simetriação de $D_{\text{KL}}(\cdot \parallel \cdot)$ (como uma *medida de distância*): Dada duas variáveis aleatórias com o mesmo suporte, a distância de Kullback-Leibler pode ser definida por

$$d_{\text{KL}}(Z, Y) = \frac{1}{2} [D_{\text{KL}}(Z \parallel Y) + D_{\text{KL}}(Y \parallel Z)] = \frac{1}{2} \int_{\mathcal{X}} [f_Z(z; \boldsymbol{\theta}_z) - f_Y(z; \boldsymbol{\theta}_y)] \log \frac{f_Z(z; \boldsymbol{\theta}_z)}{f_Y(z; \boldsymbol{\theta}_y)} dz. \quad (3.2)$$

Assim, as medidas de dissimilaridade estocástica podem ser usadas como ferramentas de processamento de imagens, uma vez que são capazes de analisar a diferença entre as distribuições que descrevem as diferentes áreas de uma imagem (FRERY *et al.*, 2014). Com o objetivo de estudar a distribuição assintótica de $d_{\text{KL}}(Z, Y)$, $D_{\text{KL}}(Z \parallel Y)$, e $D_{\text{KL}}(Y \parallel Z)$, Salicrú *et al.* (1994) propuseram uma classe de divergências, chamada *divergência*-(h, ϕ). Esta classe é uma extensão da *divergência*- ϕ proposta por Csiszár (1967) e Ali & Silvey (1966). A seguir, apresenta-se um tratamento estatístico para testes de hipóteses baseados em medidas de divergência e distância de Kullback-Leibler.

Sejam Z e Y duas variáveis aleatórias definidas anteriormente. Assumindo que ambas compartilham um suporte comum $\mathcal{X} \subset \mathbb{R}$, a divergência-(h, ϕ) entre Z e Y é definida por

$$D_{\phi}^h(X \parallel Y) = h \left(\mathbb{E}_Y \left[\phi \left(\frac{f_Z(Y; \boldsymbol{\theta}_z)}{f_Y(Y; \boldsymbol{\theta}_y)} \right) \right] \right) = h \left(\int_{\mathcal{X}} \phi \left(\frac{f_Z(x; \boldsymbol{\theta}_z)}{f_Y(x; \boldsymbol{\theta}_y)} \right) f_Y(x; \boldsymbol{\theta}_y) dx \right), \quad (3.3)$$

em que $\phi: (0, \infty) \rightarrow [0, \infty)$ é uma função convexa, $h: (0, \infty) \rightarrow [0, \infty)$ é uma função estritamente crescente com $h(0) = 0$, e formas indeterminadas são atribuídas valor zero. As medidas de divergência D_{KL} e distância d_{KL} são obtidas através de uma escolha minuciosa das funções $[h(x), \phi(x)] = [x, x \log(x)]$ e $[h(x), \phi(x)] = [1/2 x, (x - 1) \log(x)]$. Salicrú *et al.* (1994) propuseram resultados para descrever a distribuição assintótica das divergências (h, ϕ) na forma do lema abaixo.

Lema 2. *Assuma as condições de regularidade propostas por Salicrú et al. (1994, p. 380) satisfeitas. Se $\frac{m}{m+n} \xrightarrow{m, n \rightarrow \infty} \lambda \in (0, 1)$ e $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$, então*

$$\frac{2mn}{m+n} \frac{d_{\phi}^h(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2)}{h'(0)\phi''(1)} \xrightarrow{m, n \rightarrow \infty} \chi_M^2, \quad (3.4)$$

em que “ $\xrightarrow{\mathcal{D}}$ ” denota convergência em distribuição, $\widehat{\boldsymbol{\theta}}_1$ e $\widehat{\boldsymbol{\theta}}_2$ são os estimadores de máxima verossimilhança de θ_1 e θ_2 baseados em amostras independentes de tamanhos n e m , respectivamente.

Baseado no Lema 2, testes de hipóteses estatísticos para a hipótese nula $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$

podem ser derivados. Em particular, a seguinte estatística é considerada:

$$S_\phi^h(\hat{\theta}_1, \hat{\theta}_2) = \frac{2mnv}{m+n} d_\phi^h(\hat{\theta}_1, \hat{\theta}_2),$$

em que $v = 1/(h'(0)\phi''(1))$ é uma constante que depende da distância escolhida. Para as medidas D_{KL} e d_{KL} , v assume valor $1/2$ e 1 , respectivamente. Assim, um teste de hipótese geral é dado na forma da proposição subsequente.

Proposição 1. *Sejam m e n assumindo valores grandes e $S_\phi^h(\hat{\theta}_1, \hat{\theta}_2) = s$, então a hipótese nula $\theta_1 = \theta_2$ pode ser rejeitada ao nível η se $\Pr(\chi_{M}^2 > s) \leq \eta$.*

3.1 Novas medidas de contraste com base na distribuição $G\Gamma$

Nesta seção, derivamos três medidas de contraste: duas envolvendo retornos Γ e $G\Gamma$ e uma entre distribuições $G\Gamma$. Estas expressões podem ser empregadas em vários campos, tais como medidas de complexidade estatística (FRERY *et al.*, 2013a), detecção de mudanças e bordas (NASCIMENTO *et al.*, 2014), segmentação (BEAULIEU; TOUZI, 2004), classificação (KERSTEN *et al.*, 2005) e como testes de hipóteses (NASCIMENTO *et al.*, 2010). Nesta dissertação, usamos as medidas derivadas no último sentido. A seguir, apresentamos uma breve descrição e apresentação de cada fórmula.

- (a) Medida KL por assumir que o modelo é $Y \sim \Gamma$ quando os dados seguem a distribuição $Z \sim G\Gamma$:

$$\begin{aligned} D_{KL}(Z||Y) &= D_{KL}\left(G\Gamma(\theta, \mu L^{-1/\theta}, L) \parallel \Gamma(L_0, \mu_0^{-1}L_0)\right) = \log\left(|\theta| \frac{L^L}{L_0^{L_0}} \frac{\Gamma(L_0)}{\Gamma(L)} \frac{\mu_0^{L_0}}{(\mu^L)^\theta}\right) \\ &\quad + (\theta L - L_0) \left\{ \theta^{-1} [\psi(L) - \log L] + \log \mu \right\} \\ &\quad + \frac{L_0}{L^{1/\theta}} \frac{\mu}{\mu_0} \frac{\Gamma(L + \theta^{-1})}{\Gamma(L)} - L \end{aligned}$$

- (b) Medida KL por assumir que o modelo é $Z \sim G\Gamma$ quando os dados seguem a distribuição $Y \sim \Gamma$:

$$\begin{aligned} D_{KL}(Y||Z) &= D_{KL}\left(\Gamma(L_0, \mu_0^{-1}L_0) \parallel G\Gamma(\theta, \mu L^{-1/\theta}, L)\right) = \log\left(\frac{1}{|\theta|} \frac{L_0^{L_0}}{L^L} \frac{\Gamma(L)}{\Gamma(L_0)} \frac{(\mu^L)^\theta}{\mu_0^{L_0}}\right) \\ &\quad + (L_0 - \theta L) \left\{ \psi(L_0) - \log L_0 + \log \mu_0 \right\} \\ &\quad + \frac{L}{L_0^\theta} \left(\frac{\mu_0}{\mu}\right)^\theta \frac{\Gamma(L_0 + \theta)}{\Gamma(L_0)} - L_0 \end{aligned}$$

- (c) Medida KL simetrizada [conhecida como *divergência de Jensen-Shannon* (MENÉNDEZ *et al.*, 1997)] para quantificar a distância entre os modelos Γ e $G\Gamma$:

$$\begin{aligned} 2 d_{\text{KL}}(Z, Y) &= D_{\text{KL}}(Z||Y) + D_{\text{KL}}(Y||Z) = d_{\text{KL}}\left(\Gamma(L_0, \mu_0^{-1}L_0) G\Gamma(\theta, \mu L^{-1/\theta}, L)\right) \\ &= (\theta L - L_0) \left\{ \theta^{-1} [\psi(L) - \log L] + \log \mu \right\} + \frac{L_0}{L^{1/\theta}} \frac{\mu}{\mu_0} \frac{\Gamma(L + \theta^{-1})}{\Gamma(L)} \\ &\quad - L + (L_0 - \theta L) \left\{ \psi(L_0) - \log L_0 + \log \mu_0 \right\} \\ &\quad + \frac{L}{L_0^\theta} \left(\frac{\mu_0}{\mu}\right)^\theta \frac{\Gamma(L_0 + \theta)}{\Gamma(L_0)} - L_0 \end{aligned}$$

- (d) Medida de contraste KL para quantificar a distância entre dois elementos no modelo $G\Gamma$ (ou, como caso particular, dentro da distribuição Γ) na descrição de dados contaminados: Seja $Z_i \sim G\Gamma(\theta_i, \mu_i L_i^{-1/\theta_i}, L_i)$ e $Y_i \sim \Gamma(L_i, \mu_i^{-1}L_i)$, para $i = 1, 2$:

$$\begin{aligned} 2 d_{\text{KL}}(Z_1, Z_2) &= (\theta_1 L_1 - \theta_2 L_2) \left\{ \frac{\psi(L_1) - \log L_1}{\theta_1} - \left[\frac{\psi(L_2) - \log L_2}{\theta_2} \right] \right. \\ &\quad \left. + \log \mu_1 - \log \mu_2 \right\} + L_2 \left\{ \left(\frac{\mu_1}{\mu_2}\right)^{\theta_2} L_1^{-\theta_2/\theta_1} \frac{\Gamma(L_1 + \theta_2/\theta_1)}{\Gamma(L_1)} - 1 \right\} \\ &\quad + L_1 \left\{ \left(\frac{\mu_2}{\mu_1}\right)^{\theta_1} L_2^{-\theta_1/\theta_2} \frac{\Gamma(L_2 + \theta_1/\theta_2)}{\Gamma(L_2)} - 1 \right\} \end{aligned}$$

$$\xrightarrow{\theta_1=\theta_2=1}$$

$$\begin{aligned} 2 d_{\text{KL}}(Y_1, Y_2) &= (L_1 - L_2) \{ \psi(L_1) - \psi(L_2) + \log L_2 - \log L_1 + \log \mu_1 - \log \mu_2 \} \\ &\quad + L_2 \left\{ \frac{\mu_1}{\mu_2} - 1 \right\} + L_1 \left\{ \frac{\mu_2}{\mu_1} - 1 \right\}. \end{aligned}$$

Aplicando as cinco medidas derivadas na Proposição 1, cinco testes de hipóteses (NASCIMENTO *et al.*, 2010; FRERY *et al.*, 2013b; FRERY *et al.*, 2014) são derivados tendo como estatísticas as quantidades

1. $S_{\text{DZY}} = h_{m,n} D_{\text{KL}}(Z || Y)$,
2. $S_{\text{DYZ}} = h_{m,n} D_{\text{KL}}(Y || Z)$,
3. $S_{\text{dYZ}} = 2 h_{m,n} d_{\text{KL}}(Y, Z)$,
4. $S_{\text{dZZ}} = 2 h_{m,n} d_{\text{KL}}(Z_1, Z_2)$,

$$5. S_{\text{dYY}} = 2 h_{m,n} d_{\text{KL}}(Y_1, Y_2),$$

em que $h_{m,n} = nm/(n+m)$. Da Proposição 1, para amostras suficientemente grandes, as estatísticas resultantes seguem uma distribuição qui-quadrado três graus de liberdade.

3.2 Nova medida de contraste entre distribuições \mathcal{W}_3^{C} e $\text{G}\Gamma(\Gamma)$

Nesta seção, derivamos uma nova medida de contraste baseada em retornos seguindo as distribuições WCE e $\text{G}\Gamma$ trivariada não-correlacionada (denotada por $\text{G}\Gamma_3^0$). Note que até a seção anterior, apresentam-se ferramentas para identificar contraste e modelagem de normas quadradas de canais complexos (a saber HH, HV e VV). Essas divergências quantificam quaisquer modelos diferentes nas famílias Γ ou $\text{G}\Gamma$ ou entre elas. Neste ponto, estamos interessados em checar se a seguinte hipótese pode ser verificada:

$$\mathcal{H}_0 : \Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2) \equiv \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix},$$

em que $\text{diag}(\cdot)$ representa a matriz diagonal de um argumento vetorial. Para esse fim, derivamos a seguinte divergência de Kullback-Leibler por assumir que os retornos seguem o modelo $\text{G}\Gamma_3^0$ quando sua distribuição real é WCE. Sejam Z_{HH} , Z_{HV} e Z_{VV} as intensidades devido aos canais HH, HV e VV, respectivamente. Então, temos:

$$\begin{aligned}
 & D_{\text{KL}}(\mathbf{Z} \parallel \text{diag}(Z_{\text{HH}}, Z_{\text{HV}}, Z_{\text{VV}})) \equiv \\
 & D_{\text{KL}} \left(\mathcal{W}(\boldsymbol{\Sigma}, L) \parallel \begin{bmatrix} G\Gamma(\theta_{\text{HH}}, L_{\text{HH}}, \mu_{\text{HH}}) & 0 & 0 \\ 0 & G\Gamma(\theta_{\text{HV}}, L_{\text{HV}}, \mu_{\text{HV}}) & 0 \\ 0 & 0 & G\Gamma(\theta_{\text{VV}}, L_{\text{VV}}, \mu_{\text{VV}}) \end{bmatrix} \right) \\
 & = (L - m)[\log |\boldsymbol{\Sigma}| + \psi_m^{(0)}(L) - m \log L] \\
 & \quad - Lm - (\theta_{\text{HH}}L_{\text{HH}} - 1)[\psi^{(0)}(L_{\text{HH}}) - \log L_{\text{HH}} + \log(2\mu_{\text{HH}})] \\
 & \quad + \left[\frac{L_{\text{HH}}}{(2\mu_{\text{HH}})^{\theta_{\text{HH}}}} \right] \left\{ \frac{L_{\text{HH}}^{L_{\text{HH}}}}{(2\mu_{\text{HH}})^{L_{\text{HH}}}\Gamma(L_{\text{HH}})} \frac{\Gamma(L_{\text{HH}}\theta_{\text{HH}} + L_{\text{HH}} - 1)}{\left[\frac{L_{\text{HH}}}{2\mu_{\text{HH}}} \right]^{(L_{\text{HH}}\theta_{\text{HH}} + L_{\text{HH}} - 1)}} \right\} \\
 & \quad - (\theta_{\text{HV}}L_{\text{HV}} - 1)[\psi^{(0)}(L_{\text{HV}}) - \log L_{\text{HV}} + \log(2\mu_{\text{HV}})] \\
 & \quad + \left[\frac{L_{\text{HV}}}{(2\mu_{\text{HV}})^{\theta_{\text{HV}}}} \right] \left\{ \frac{L_{\text{HV}}^{L_{\text{HV}}}}{(2\mu_{\text{HV}})^{L_{\text{HV}}}\Gamma(L_{\text{HV}})} \frac{\Gamma(L_{\text{HV}}\theta_{\text{HV}} + L_{\text{HV}} - 1)}{\left[\frac{L_{\text{HV}}}{2\mu_{\text{HV}}} \right]^{(L_{\text{HV}}\theta_{\text{HV}} + L_{\text{HV}} - 1)}} \right\} \\
 & \quad - (\theta_{\text{VV}}L_{\text{VV}} - 1)[\psi^{(0)}(L_{\text{VV}}) - \log L_{\text{VV}} + \log(2\mu_{\text{VV}})] \\
 & \quad + \left[\frac{L_{\text{VV}}}{(2\mu_{\text{VV}})^{\theta_{\text{VV}}}} \right] \left\{ \frac{L_{\text{VV}}^{L_{\text{VV}}}}{(2\mu_{\text{VV}})^{L_{\text{VV}}}\Gamma(L_{\text{VV}})} \frac{\Gamma(L_{\text{VV}}\theta_{\text{VV}} + L_{\text{VV}} - 1)}{\left[\frac{L_{\text{VV}}}{2\mu_{\text{VV}}} \right]^{(L_{\text{VV}}\theta_{\text{VV}} + L_{\text{VV}} - 1)}} \right\}.
 \end{aligned}$$

Note que $D_{\text{KL}}(\mathbf{Z} \parallel \text{diag}(Z_{\text{HH}}, Z_{\text{HV}}, Z_{\text{VV}}))$ representa uma informação “quase mútua” para o modelo Wishart complexo escalonado (COVER; THOMAS, 1991). Como discutido por Salicrú *et al.* (1994), sob a hipótese nula e considerando L de \mathbf{Z} e L_M de Z_M para $M \in \{\text{HH}, \text{HV}, \text{VV}\}$ como parâmetros de incômodo (ou conhecidos e obtidos a partir do processamento PolSAR), a estatística

$$S_{\mathbf{Z}\mathbf{Z}} = h_{m,n} D_{\text{KL}}(\mathbf{Z} \parallel \text{diag}(Z_{\text{HH}}, Z_{\text{HV}}, Z_{\text{VV}}))$$

segue um modelo χ^2 para grandes amostras. De acordo com Kupperman (1957), considerando apenas uma amostra, o teste estatístico para a hipótese nula $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ baseado na divergência de Kullback-Leibler tem o seguinte comportamento assintótico

$$2n D_{\text{KL}}(\hat{\boldsymbol{\theta}} \parallel \boldsymbol{\theta}_0) = 2n \mathbb{E}_{\hat{\boldsymbol{\theta}}} \left[\log \frac{f_{\hat{\boldsymbol{\theta}}}(Z)}{f_{\boldsymbol{\theta}_0}(Z)} \right] = 2n \int_{\mathcal{X}} \log \frac{f_{\hat{\boldsymbol{\theta}}}(z)}{f_{\boldsymbol{\theta}_0}(z)} dF_{\hat{\boldsymbol{\theta}}}(z) dz \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \chi_M^2, \quad (3.5)$$

em que \mathcal{X} é o suporte da variável aleatória Z em $(\sigma_{\mathcal{X}}, \text{Pr}_{\boldsymbol{\theta}})$ com função de distribuição acumulada $F_{\boldsymbol{\theta}}(z)$, M é a dimensão de $\boldsymbol{\theta} = (\theta_1, \dots, \theta_M)^\top$ e $\hat{\boldsymbol{\theta}}$ é o EMV para $\boldsymbol{\theta}$.

Neste capítulo, quantifica-se o desempenho das três propostas desta dissertação:

- (i) Estimadores para os parâmetros do modelo GF: número de *looks* e forma por exponencialização. Adicionalmente, propõe-se um estimador semi-paramétrico para o número de *looks* com desempenho comparável à proposta do Anfinsen *et al.* (2009);
- (ii) Estimadores melhorados segundo a abordagem de Cox-Snell para os parâmetros da distribuição GF;
- (iii) Medidas de divergência em/entre distribuições GF e WCE. Esta última podendo ser utilizada como uma medida de redundância na suposição de que os retornos PolSAR *multilook* são WCE distribuídos.

Inicialmente os métodos são avaliados sob dados sintéticos obedecendo um esquema de Monte Carlo. Subsequentemente, sua relevância é quantificada em cenários reais tanto em intensidades SAR como em imagens *Full PolSAR*.

Os resultados referentes a (i) são apresentados na Seção 4.1. Os resultados de (ii) estão na Seção 4.2. Finalmente, os resultados relativos a (iii) são apresentados na Seção 4.3.

4.1 Estimação por MV para os parâmetros do modelo $G\Gamma$ e estimador semi paramétrico para o número de *looks*

4.1.1 Desempenho em imagens sintéticas

Inicialmente, realiza-se um estudo de simulação em que os dados são gerados a partir da distribuição $G\Gamma$, com tamanhos de amostra, $n \in \{64, 100, 144\}$. Como cenários deste estudo de simulação, escolhemos:

- (i) $|\theta| \in \{0.8, 1.0, 4.0\}$. $|\theta| \geq 1$ representa cenários homogêneos, enquanto $|\theta| < 1$ caracteriza regiões heterogêneas;
- (ii) $L \in \{1, 4, 8\}$. O aumento de L implica em imagens resultantes menos afetadas pelo efeito do ruído *speckle*, e $L = 1$ representa o procedimento *singlelook*;
- (iii) $\mu = 1$. Assumimos esta suposição por duas razões: (a) este parâmetro pode ser entendido como uma quantidade conhecida dada pela média amostral das intensidades observadas, e (b) Freitas *et al.* (2005) adotaram $\mu = 1$ para cenários SAR polarimétricos.

Utilizam-se 10.000 replicações Monte Carlo e, em cada uma, computam-se:

- (i) Para o estimador MV de θ , $\hat{\theta}$, a média dos EMVs e o erro quadrático médio (EQM) e
- (ii) Para estimadores MV de L , $\hat{L}(\theta + \epsilon)$, e μ , $\hat{\mu}(\theta + \epsilon)$, a média dos vieses de seus EMVs para diferentes graus de contaminação $\epsilon \in \{0\%, 1\%, 2\%, 3\%, 7\%, 8\%, 9\%, 10\%\}$, quando se deseja estimar $[L, \mu] = [1, 1]$.

Tabela 4.1 apresenta os resultados do estudo para os EMVs de θ . Em geral, como esperado, tem-se indícios de que a propriedade de consistência—(i) $\lim_{n \rightarrow \infty} E(\hat{\theta}_n) = \theta$ e (ii) $\lim_{n \rightarrow \infty} \text{Var}(\hat{\theta}_n) = 0$ —ocorre quando o tamanho da amostra aumenta. Além disso, com base nos valores assumidos pelo EQM, a condição $|\theta| \rightarrow \infty$ gera cenários mais complicados.

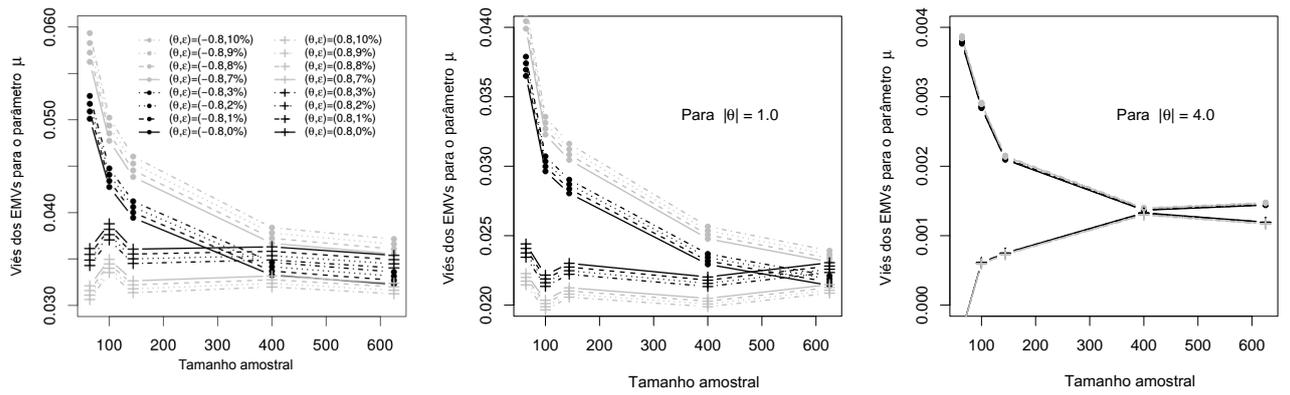
Figura 4.1 apresenta um estudo da influência do viés sobre os EMVs para μ e L . Os primeiros resultados indicam que um erro de $\epsilon\%$ no EMV de θ , $\theta(1 + \epsilon)$, tem influência menor sobre as expressões propostas, $\hat{\mu}(\cdot)$ e $\hat{L}(\cdot)$, quando $|\theta|$ decresce. Este fato pode ser explicado pela expressão do coeficiente de variação dado por Li *et al.* (2011), que é inversamente proporcional ao aumento de θ sob L fixo, ou seja, espera-se que o coeficiente de variação aumente quando a quantidade $L\theta^2$ decresce. Comparativamente, EMVs para μ são menos viesados que aqueles para L . Este último resultado indica que pode ser interessante corrigir estimadores para L , como através das expressões 2.22 e 2.24.

Note que 2.10 é uma expressão para estimação do NEL com base nas informações de apenas um canal. Na prática, temos a informação de três (a saber HH, HV e VV). A fim de usar as informações dos três, propôs-se o estimador semi-paramétrico 2.14.

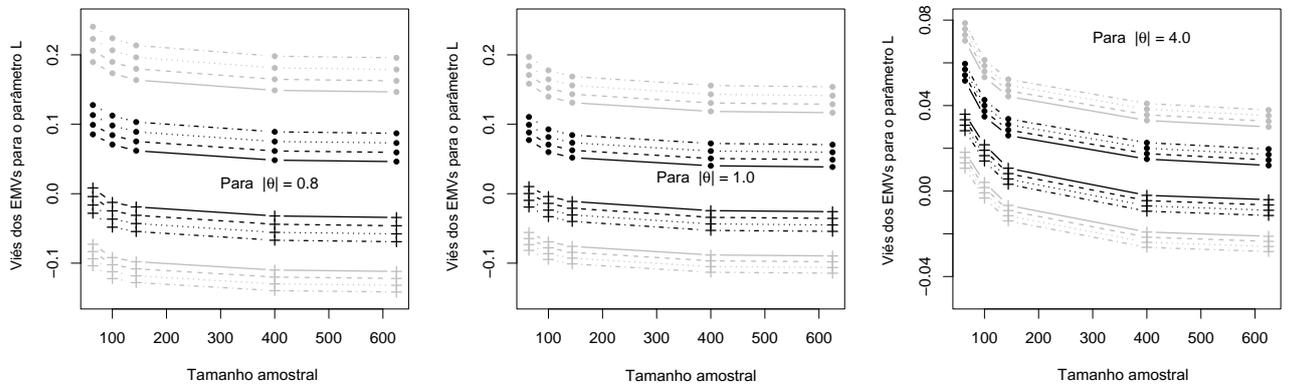
Tabela 4.2 apresenta a média das estimativas e o EQM do estimador 2.14 comparativamente à proposta de Anfinsen *et al.* (2009) dada em 2.13. Nota-se que o estimador semi paramétrico funciona melhor para menores números de *looks* e tamanhos de amostra. Convém ressaltar que nas regiões oceânica e urbana com $L = 8$, o estimador proposto por Anfinsen *et al.* (2009) se sobressai.

Tabela 4.1: Performance dos EMVs para o parâmetro θ do modelo GF

n	θ	$\widehat{\theta} (\text{EQM}(\widehat{\theta}))$					
		$L = 1$		$L = 4$		$L = 8$	
64	-4.0	-4.522	(5.726)	-4.679	(8.641)	-5.241	(12.810)
		-4.348	(3.079)	-4.330	(5.218)	-4.686	(8.271)
		-4.216	(1.781)	-4.150	(3.465)	-4.351	(5.635)
100	-1.0	-1.135	(0.584)	-1.175	(0.652)	-1.335	(1.066)
		-1.096	(0.216)	-1.086	(0.335)	-1.169	(0.537)
		-1.051	(0.112)	-1.052	(0.224)	-1.095	(0.367)
144	-0.8	-0.878	(0.464)	-0.956	(0.488)	-1.079	(0.729)
		-0.868	(0.154)	-0.871	(0.218)	-0.931	(0.340)
		-0.842	(0.069)	-0.828	(0.134)	-0.876	(0.229)
	0.8	0.947	(0.483)	0.963	(0.488)	1.082	(0.685)
		0.870	(0.140)	0.865	(0.217)	0.934	(0.347)
		0.842	(0.070)	0.826	(0.142)	0.868	(0.226)
	1.0	1.180	(0.667)	1.197	(0.734)	1.343	(1.065)
		1.088	(0.207)	1.072	(0.347)	1.160	(0.534)
		1.056	(0.113)	1.042	(0.222)	1.094	(0.361)
	4.0	4.568	(5.860)	4.696	(8.831)	5.238	(13.308)
		4.374	(3.153)	4.329	(5.088)	4.666	(8.331)
		4.202	(1.779)	4.148	(3.472)	4.387	(5.625)



(a) Média de Monte Carlo para o viés de $\hat{\mu}(\theta + \epsilon)$ para $\theta = \pm 0.8, \pm 1.0, \pm 4.0$ e $\epsilon = 0\%, 1\%, 2\%, 3\%, 7\%, 8\%, 9\%, 10\%$.



(b) Média de Monte Carlo para o viés de $\hat{L}(\theta + \epsilon)$ para $\theta = \pm 0.8, \pm 1.0, \pm 4.0$ e $\epsilon = 0\%, 1\%, 2\%, 3\%, 7\%, 8\%, 9\%, 10\%$.

Figura 4.1: Estudo das propriedades assintóticas dos vieses de $\hat{\mu}(\theta)$ e $\hat{L}(\theta)$ sob $\mu = 1$, $L = 1$ e $|\theta| \in \{0.8, 1.0, 4.0\}$.

Tabela 4.2: Estimativas MV para L e seus respectivos EQMs para dados SAR sintéticos

Regiões	L	n	L (GF $L=4$)	EQM	L (Anfinsen)	EQM
Oceano	3	64	3.0056	0.0241	2.3504	1.3676
		100	3.0016	0.0129	2.7761	0.4626
		144	3.0012	0.0073	2.4302	1.0728
	4	64	4.0020	0.0319	4.1074	0.0532
		100	4.0023	0.0175	4.0338	0.0217
		144	4.0016	0.0100	4.0168	0.0161
	8	64	8.0063	0.0677	8.1429	0.0654
		100	8.0009	0.0363	8.0492	0.0191
		144	8.0023	0.0204	8.0248	0.0091
Floresta	3	64	3.0066	0.0108	2.3568	1.3566
		100	3.0016	0.0055	2.7668	0.4791
		144	3.0008	0.0031	2.4180	1.0946
	4	64	4.0051	0.0142	4.1055	0.0571
		100	4.0028	0.0075	4.0347	0.0227
		144	4.0010	0.0041	4.0181	0.0125
	8	64	8.0048	0.0293	8.1430	0.0650
		100	8.0012	0.0153	8.0495	0.0191
		144	8.0011	0.0087	8.0249	0.0091
Urbana	3	64	3.0065	0.0144	2.3593	1.3549
		100	3.0020	0.0074	2.7766	0.4609
		144	3.0009	0.0042	2.4231	1.0863
	4	64	4.0049	0.0195	4.1059	0.0583
		100	4.0041	0.0102	4.0334	0.0253
		144	4.0017	0.0058	4.0178	0.0143
	8	64	8.0062	0.0399	8.1427	0.0651
		100	8.0004	0.0214	8.0493	0.0190
		144	8.0022	0.0118	8.0251	0.0090

4.1.2 Aplicação com dados SAR de intensidade

Adicionalmente ao estudo de sintético prévio, realiza-se um experimento a dados reais com intensidades SAR. Inicialmente pretende-se destacar o uso do método de estimação proposto considerando o número de *looks* para os modelos Γ obtidos por meio dos processos (conhecido) e desconhecido. Paralelamente, os resultados são comparados com a distribuição $\Gamma(L, L\mu^{-1})$ com fdp dada pela Equação (2.2). Para o modelo Γ , o EMV para μ é dado pela média amostral; enquanto, aquele para L é definido pela solução da equação não-linear-caso particular para $(\theta = 1)$ da Equação (2.11):

$$\log \widehat{L}(1) - \psi^{(0)}(\widehat{L}(1)) - \log \left(n^{-1} \sum_{i=1}^n z_i \right) + n^{-1} \sum_{i=1}^n \log z_i = 0,$$

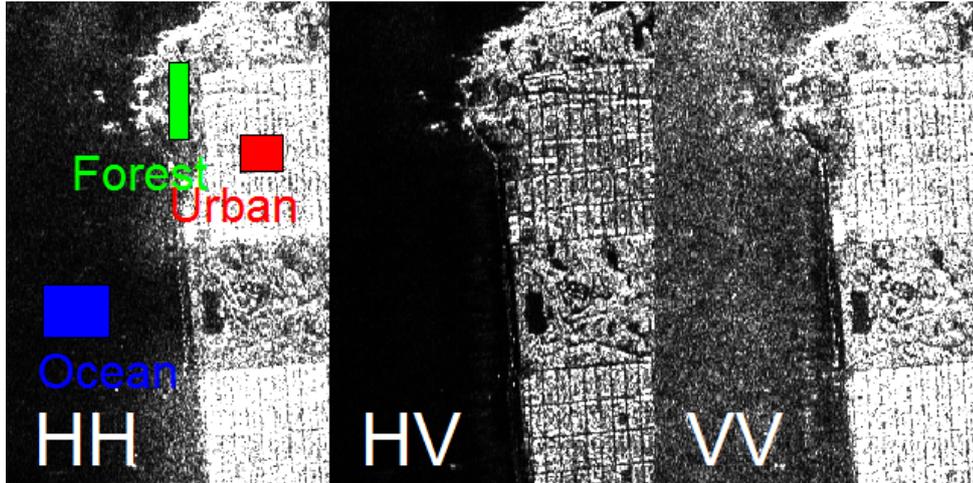
em que z_1, \dots, z_n é uma amostra observada. Para calcular os erros padrão (EPs) dos EMVs para Γ , pode-se considerar a matriz de informação de Fisher proposta com θ conhecido e igual a um.

Em particular, consideramos a imagem apresentada na Figura 1.2(b). A Figura 4.2(a) ilustra a imagem SAR nos três canais de polarização HH, HV e VV e algumas regiões selecionadas. Tabela 4.3 apresenta uma análise descritiva (\bar{z} : média amostral, \tilde{z} : mediana amostral, $CV(z)$: Coeficiente de variação, $z(1)$: mínimo, $z(n)$: máximo, γ_1 : assimetria, γ_2 : curtose e o tamanho amostral) para os dados obtidos:

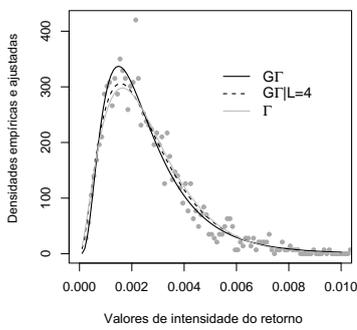
- a) a diferença $|\bar{Z} - \tilde{Z}|$ assume valores mais altos para regiões urbanas,
- b) valores CV maiores (que um) indicam texturas pronunciadas (como as apresentadas por regiões urbanas),
- c) em todos os casos (baseado em $\gamma_1 > 0$), valores superiores estão concentrados à esquerda da média, com valores extremos para a direita, e
- d) regiões urbanas ($\gamma_2 > 3$) são caracterizadas por distribuições leptocúrticas, que apresentam alta probabilidade para valores extremos.

Tabela 4.3: Estatísticas descritivas para os dados SAR coletados

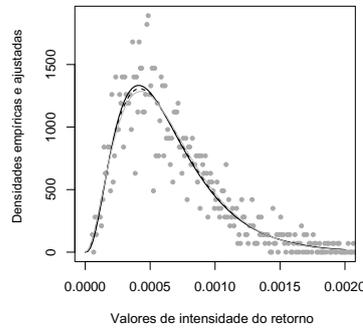
Region	Channel	Descriptive Statistics							
		\bar{z}	\tilde{z}	$CV(z)$	$z(1)$	$z(n)$	γ_1	γ_2	tamanho
Ocean	HH	0.0026	0.0022	71.91	0.00024	0.027	36.74	4.02	1428
	HV	0.0006	0.0005	58.40	0.00005	0.003	2.06	1.23	1428
	VV	0.0172	0.0148	64.44	0.00194	0.141	13.45	2.22	1428
Forest	HH	0.0477	0.0329	251.50	0.00222	2.779	428.05	19.42	624
	HV	0.0304	0.0208	103.40	0.00045	0.228	7.85	2.36	624
	VV	0.0510	0.0354	121.52	0.00148	1.112	138.41	9.01	624
Urban	HH	0.2185	0.1039	231.36	0.00936	9.371	180.58	11.35	621
	HV	0.0426	0.0243	184.74	0.00146	1.462	172.92	10.79	621
	VV	0.1988	0.0896	214.53	0.00583	6.048	75.60	7.46	621



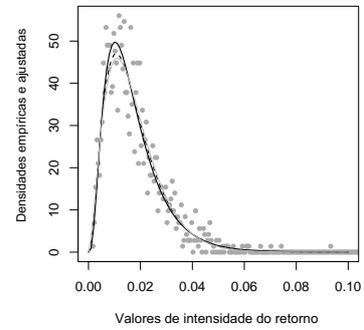
(a) São Francisco (E.U.A.)



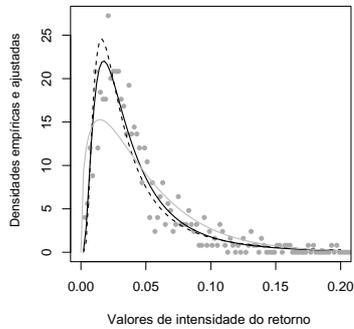
(b) Canal HH em Oceano



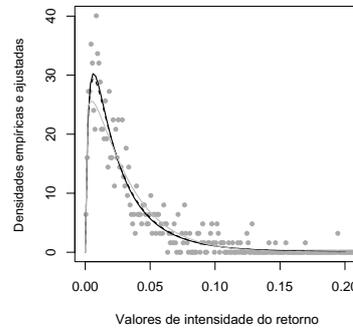
(c) Canal HV em Oceano



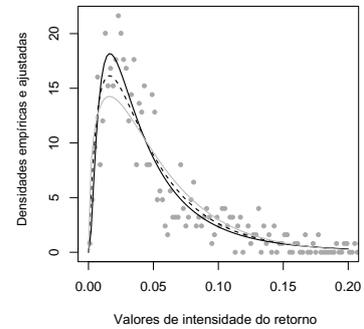
(d) Canal VV em Oceano



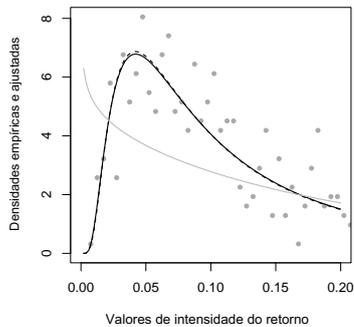
(e) Canal HH em Floresta



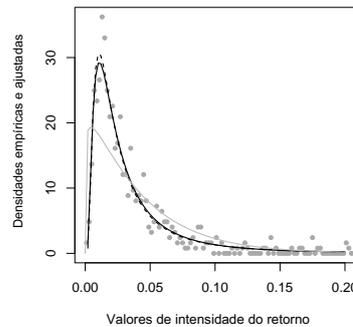
(f) Canal HV em Floresta



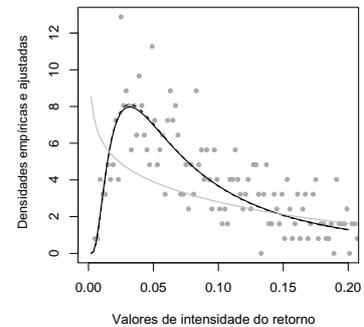
(g) Canal VV em Floresta



(h) Canal HH em Zona Urbana



(i) Canal HV em Zona Urbana



(j) Canal VV em Zona Urbana

Figura 4.2: Densidades empíricas (\bullet) vs. ajustadas para distribuições GT (curvas sólidas pretas), GT especificado $L = 4$ (curvas tracejadas pretas) e Γ (curvas sólidas cinzas).

Baseado nos dados SAR reais, ajustamos o modelo Γ sem informação adicional, o modelo Γ com informação adicional (tal como discutido no Algoritmo 1) denotado por “ $\Gamma|L = 4$ ”, a distribuição Γ e a distribuição WCE. Tabela 4.4 exhibe as estimativas de MV e seus respectivos EPs para todos os casos. Com relação às EMVs para o parâmetro θ nos modelos Γ e $\Gamma|L = 4$, observamos uma tendência: as estimativas em um mesmo canal de polarização diminuem com o aumento da intensidade da textura (i.e. Oceano \ll Floresta \ll Urbana), assim como os EPs diminuem neste mesmo sentido, com exceção do canal HH para o modelo $\Gamma|L = 4$. Ressalta-se que, as estimativas para o modelo $\Gamma|L = 4$ foram superiores quando comparados com Γ . Com respeito às EMVs para o parâmetro L , nota-se que os maiores valores de EP ocorrem sob a suposição Γ nos casos (Oceano,HH) e (Floresta,{HH,VV}). No entanto, estes valores melhoram quando se considera a informação adicional do número de *looks* na estimação de θ , isto é, o modelo $\Gamma|L = 4$. Em todos os casos, as estimativas MV para L se mostraram as menos viesadas sob este último modelo. Como L é conhecido para o processamento de imagens SAR, analisamos o viés, $B(\hat{L})$, dos EMVs:

- Para o modelo Γ , $B(\hat{L}) \in [0.665, 101.324]$;
- Para o modelo $\Gamma|L = 4$, $B(\hat{L}) \in [0.014, 0.206]$;
- Para o modelo Γ , $B(\hat{L}) \in [0.933, 3.181]$.

Tabela 4.4: EMVs e seus EPs estimados para os dados SAR de intensidade

Região	Canal	Modelo	$\hat{\delta}$ (EPs)						
Oceano	HH	$\Gamma(\theta, L, \mu)$	0.304	(6.83×10^{-2})	26.817	(12.00)	2.293×10^{-3}	(2.33×10^{-5})	
		$\Gamma(\theta, L, \mu) L = 4$	0.807	(5.34×10^{-2})	3.873	(0.49)	2.533×10^{-3}	(4.35×10^{-5})	
		$\Gamma(L, \mu)$	•		2.662	(0.47)	2.631×10^{-3}	(6.46×10^{-5})	
	HV	$\Gamma(\theta, L, \mu)$	0.729	(9.95×10^{-2})	5.527	(1.43)	6.083×10^{-4}	(1.08×10^{-5})	
		$\Gamma(\theta, L, \mu) L = 4$	0.866	(9.48×10^{-2})	3.978	(0.82)	6.224×10^{-4}	(1.21×10^{-5})	
		$\Gamma(L, \mu)$	•		3.067	(0.74)	6.356×10^{-4}	(1.62×10^{-5})	
	VV	$\Gamma(\theta, L, \mu)$	0.491	(8.03×10^{-2})	10.695	(3.41)	1.571×10^{-2}	(2.17×10^{-4})	
		$\Gamma(\theta, L, \mu) L = 4$	0.821	(7.06×10^{-2})	3.937	(0.64)	1.671×10^{-2}	(3.04×10^{-4})	
		$\Gamma(L, \mu)$	•		2.774	(0.60)	1.727×10^{-2}	(4.52×10^{-4})	
	Floresta	HH	$\Gamma(\theta, L, \mu)$	-0.125	(6.70×10^{-2})	105.324	(112.46)	3.123×10^{-2}	(3.35×10^{-4})
			$\Gamma(\theta, L, \mu) L = 4$	-0.654	(4.60×10^{-2})	3.794	(0.53)	2.658×10^{-2}	(6.28×10^{-4})
			$\Gamma(L, \mu)$	•		1.435	(0.19)	4.779×10^{-2}	(2.20×10^{-3})
HV		$\Gamma(\theta, L, \mu)$	0.413	(7.57×10^{-2})	5.828	(2.04)	2.339×10^{-2}	(7.00×10^{-4})	
		$\Gamma(\theta, L, \mu) L = 4$	0.505	(7.14×10^{-2})	3.975	(1.06)	2.445×10^{-2}	(8.07×10^{-4})	
		$\Gamma(L, \mu)$	•		1.186	(0.33)	3.046×10^{-2}	(2.34×10^{-3})	
VV		$\Gamma(\theta, L, \mu)$	0.176	(7.23×10^{-2})	42.346	(34.63)	3.739×10^{-2}	(5.78×10^{-4})	
		$\Gamma(\theta, L, \mu) L = 4$	0.586	(5.48×10^{-2})	3.860	(0.70)	4.364×10^{-2}	(1.27×10^{-3})	
		$\Gamma(L, \mu)$	•		1.463	(0.29)	5.106×10^{-2}	(2.65×10^{-3})	
Urbana		HH	$\Gamma(\theta, L, \mu)$	-0.482	(5.79×10^{-2})	4.665	(1.07)	9.017×10^{-2}	(2.13×10^{-3})
			$\Gamma(\theta, L, \mu) L = 4$	-0.524	(5.58×10^{-2})	3.986	(0.81)	8.853×10^{-2}	(2.17×10^{-3})
			$\Gamma(L, \mu)$	•		0.893	(0.20)	0.218×10^{-1}	(2.15×10^{-2})
	HV	$\Gamma(\theta, L, \mu)$	-0.403	(6.14×10^{-2})	7.742	(2.31)	2.166×10^{-2}	(4.46×10^{-4})	
		$\Gamma(\theta, L, \mu) L = 4$	-0.570	(5.33×10^{-2})	3.942	(0.71)	2.031×10^{-2}	(4.90×10^{-4})	
		$\Gamma(L, \mu)$	•		1.110	(0.23)	4.265×10^{-2}	(3.12×10^{-3})	
	VV	$\Gamma(\theta, L, \mu)$	-0.440	(5.49×10^{-2})	4.880	(1.17)	7.584×10^{-2}	(1.85×10^{-3})	
		$\Gamma(\theta, L, \mu) L = 4$	-0.489	(5.25×10^{-2})	3.984	(0.82)	7.396×10^{-2}	(1.90×10^{-3})	
		$\Gamma(L, \mu)$	•		0.819	(0.25)	1.989×10^{-1}	(2.82×10^{-2})	

Portanto, o estimador $\hat{L}(\hat{\theta})$ tal que $\hat{\theta}$ é obtido assumindo uma informação adicional para L apresentou o melhor desempenho. Propostas de estimadores para o número de *looks* (co-

nhecido como número equivalente de *looks*) tem sido estudadas na literatura, a exemplo de Anfinson *et al.* (2009). Resultados fornecem evidências de que a expressão em forma fechada para \hat{L} pode consistir em uma importante ferramenta para o processamento de imagens SAR. Com relação às EMVs para os parâmetros da $G\Gamma$, aquelas para μ apresentaram os menores EPs em todos os casos, tais como indicado no estudo de simulação.

Uma vez que o modelo Γ está encaixado na distribuição $G\Gamma$, o teste da razão entre verossimilhança (TRV) é adequado para testar $\mathcal{H}_0: G\Gamma = \Gamma$. Em particular, sob \mathcal{H}_0 , $S_{RV} \sim \chi_1^2$ pode ser entendido como uma medida de distância entre os modelos irrestrito, $G\Gamma$, e restrito, Γ . A Tabela 4.5 apresenta valores de S_{LR} e seus respectivos *p-valor*. Pode-se verificar que os dados fornecem evidência da existência de uma diferença significativa entre os modelos Γ e $G\Gamma$ (ou, equivalentemente, que $\theta \neq 1$) em todos os cenários considerados. Adicionalmente, percebe-se que os valores S_{LR} aumentam conforme a intensificação da textura; isto é, de acordo com a sequência Oceano \ll Floresta \ll Urbana. Como esperado, este fato indica a existência de distâncias mais significativas nas áreas urbanas; enquanto que em regiões de oceano, S_{LR} assume os menores valores. No caso particular {Oceano, HV}, para um nível nominal de 5%, não existe diferença na escolha entre os modelos Γ e $G\Gamma$. Estes resultados foram apontados pelo EMV para θ sob a suposição $G\Gamma(\theta, L, \mu)|L = 4$, que assumiu o valor mais próximo para $\theta = 1$ (i.e., $\Gamma \equiv G\Gamma$).

Tabela 4.5: Estatísticas RV e seus p-valor sob $\mathcal{H}_0: G\Gamma = \Gamma$

Hipótese nula	Região	Canal	S_{RV}	p-valor
$\mathcal{H}_0: \theta = 1$ ($\Gamma \times G\Gamma$)	Oceano	HH	55,786	0,000
		HV	4,906	0,027
		VV	21,580	0,000
	Floresta	HH	227,024	0,000
		HV	24,387	0,000
		VV	63,746	0,000
	Urbana	HH	294,237	0,000
		HV	228,042	0,000
		VV	281,077	0,000

Com o propósito de verificar se as conclusões obtidas na Tabela 4.5, são também constatadas em toda imagem SAR, mapas de *p-valor* para os canais HH, HV e VV são elaborados na Figura 4.3. Com base neles, observa-se que os *p-valor*s diminuem-ou seja, as diferenças entre a log-verossimilhança $G\Gamma$ e Γ aumentam-quando as texturas se tornam mais pronunciadas. Em outras palavras, maiores diferenças são obtidas em cenários urbanos.

Para quantificar o desempenho dos modelos em dados SAR de intensidade, consideramos os valores de três medidas de bondade de ajuste (*goodness-of-fit-GoF*) como critério de comparação: Critério de informação de Akaike (*Akaike information criterion-AIC*), sua versão corrigida (*AICc*) e Critério de informação bayesiano (*Bayesian information criterion-BIC*). Em uma classe de modelos encaixados, vale-se a regra “menores valores

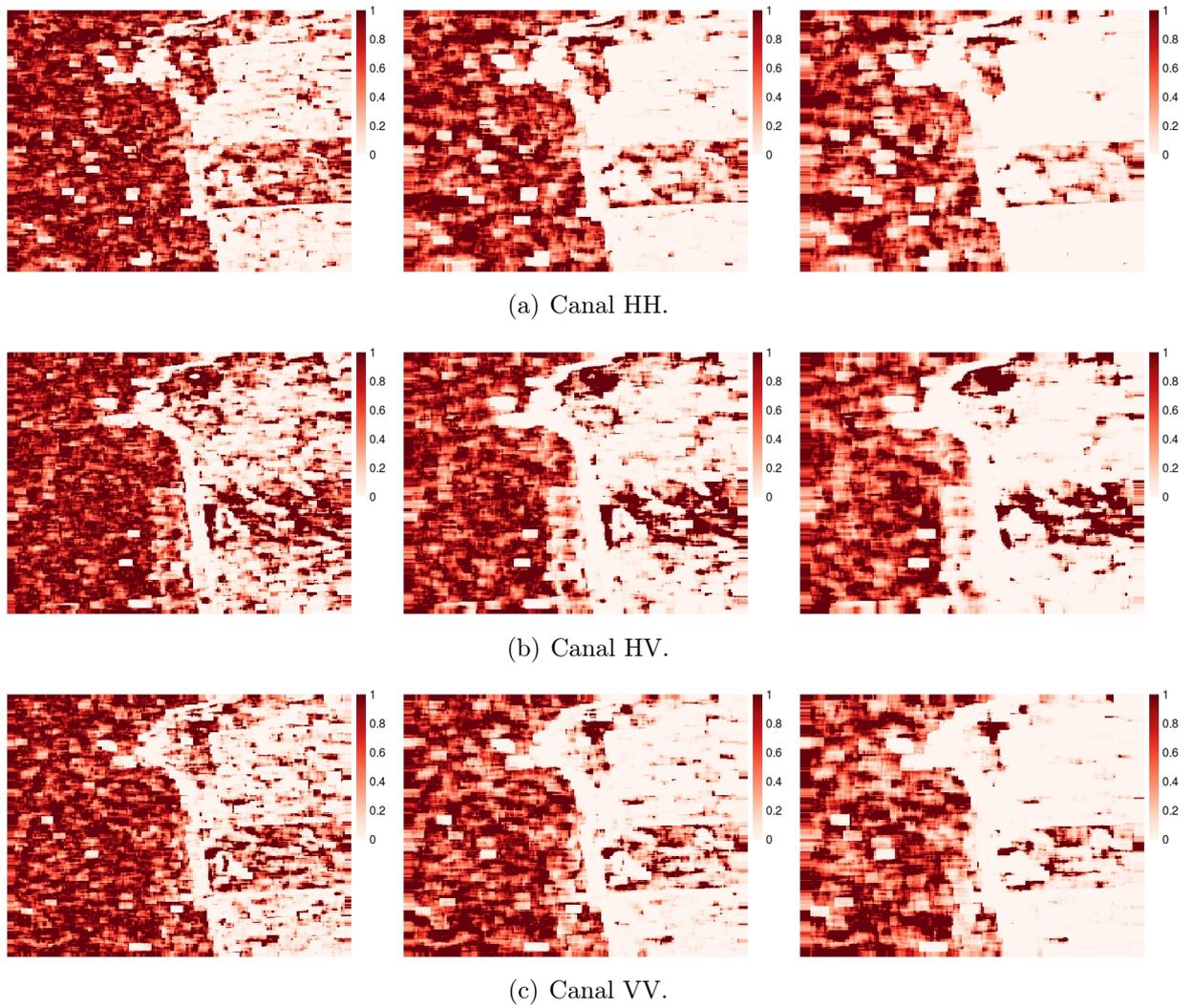


Figura 4.3: p -valores para os testes da razão entre verossimilhanças sob $\mathcal{H}_0: G\Gamma = \Gamma$ e $N = 49, 81, 121$, usando janelamentos 7×7 , 9×9 e 11×11 , respectivamente.

estão associados a melhores ajustes”. Tabela 4.6 apresenta os EMVs e valores das medidas GoF consideradas baseada em dados reais. Com exceção da medida BIC no caso {Oceano, HV}, os resultados evidenciam que os melhores desempenhos foram apresentados pelo modelo $G\Gamma$, sendo a diferença mais expressiva na região Urbana. Tabela 4.7

Tabela 4.6: Medidas de bondade de ajuste para os modelos $G\Gamma$, $G\Gamma|L = 4$ e Γ nos canais HH, HV e VV

Regiões	Medidas	Para o canal HH			Para o canal HV			Para o canal VV		
		$G\Gamma$	$G\Gamma L = 4$	Γ	$G\Gamma$	$G\Gamma L = 4$	Γ	$G\Gamma$	$G\Gamma L = 4$	Γ
Oceano	AIC	-14,753.468	-14,723.568	-14,699.683	-18,905.621	-18,904.310	-18,902.715	-9,387.137	-9,377.706	-9,367.557
	AICc	-14,753.451	-14,723.552	-14,699.674	-18,905.604	-18,904.293	-18,902.706	-9,387.120	-9,377.689	-9,367.549
	BIC	-14,737.676	-14,707.776	-14,689.154	-18,889.828	-18,888.518	-18,892.186	-9,371.345	-9,361.914	-9,357.029
Floresta	AIC	-2,813.007	-2783.171	-2,587.983	-3,138.371	-3,137.640	-3,115.985	-2,571.628	-2,555.369	-2,509.883
	AICc	-2,812.968	-2783.133	-2,587.964	-3,138.333	-3,137.601	-3,115.965	-2,571.590	-2,555.331	-2,509.864
	BIC	-2,799.698	-2769.863	-2579.111	-3,125.063	-3,124.331	-3,107.112	-2,558.320	-2,542.061	-2,501.011
Urbana	AIC	-940.7515	-940.5751	-648.5143	-2902.371	-2899.606	-2676.329	-1056.766	-1056.494	-777.6886
	AICc	-940.7126	-940.5362	-648.4949	-2902.332	-2899.568	-2676.309	-1056.727	-1056.455	-777.6692
	BIC	-927.4575	-927.2811	-639.6516	-2889.077	-2886.312	-2667.466	-1043.472	-1043.2	-768.8259

apresenta a média das estimativas e EQMs para L obtidos a partir do nosso método (considerando todos os canais) como também através do método derivado por Anfinsen *et al.* (2009). Conforme pode ser verificado, as estimativas obtidas por meio do estimador semi-paramétrico funcionam bem para pequenas amostras, sendo menos tendenciosas e mais consistentes do que aquelas obtidas através do estimador proposto em Anfinsen *et al.* (2009).

Tabela 4.7: Estimativas MV para L e seus respectivos EQMs para dados SAR reais

Regiões	n	L ($G\Gamma L = 4$)	EQM	L (Anfinsen)	EQM
Oceano	64	4.0020	0.0320	4.1074	0.0532
	100	4.0023	0.0175	4.0338	0.0217
	144	4.0016	0.0100	4.0168	0.0162
Floresta	64	4.0051	0.0142	4.1055	0.0571
	100	4.0028	0.0075	4.0347	0.0227
	144	4.0010	0.0041	4.0181	0.0125
Urbana	64	4.0049	0.0195	4.1059	0.0583
	100	4.0041	0.0102	4.0334	0.0253
	144	4.0017	0.0058	4.0179	0.0143

Figuras 4.4- 4.5 indicam que melhores estimativas são obtidas sob a suposição $G\Gamma|L = 4$. Conforme pode ser observado, Figura 4.5 apresenta um tom mais escuro (estimativas próximas do número real de *looks*, 4) que Figura 4.4. Ressalta-se que as estimativas melhoram quando N aumenta. Adicionalmente o modelo Γ apresenta um bom desempenho na estimação de L para a região de oceano ($\theta = 1$), que diminui à medida que as regiões apresentam texturas mais pronunciadas.

Figura 4.6 aponta que melhores estimativas para L , em termos de EQM, são obtidas por meio do estimador semi-paramétrico proposto nesta dissertação, quando comparadas a proposta de Anfinsen *et al.* (2009). Verifica-se também que as amplitudes das escalas de variação para as estimativas de L obtidas pelo estimador semi-paramétrico são sempre menores e possuem menor desvio em relação ao valor real de *looks*, comparativamente a proposta de Anfinsen *et al.* (2009). Destaca-se que esta última funciona bem em regiões de oceano, em termos de estimação do NEL.

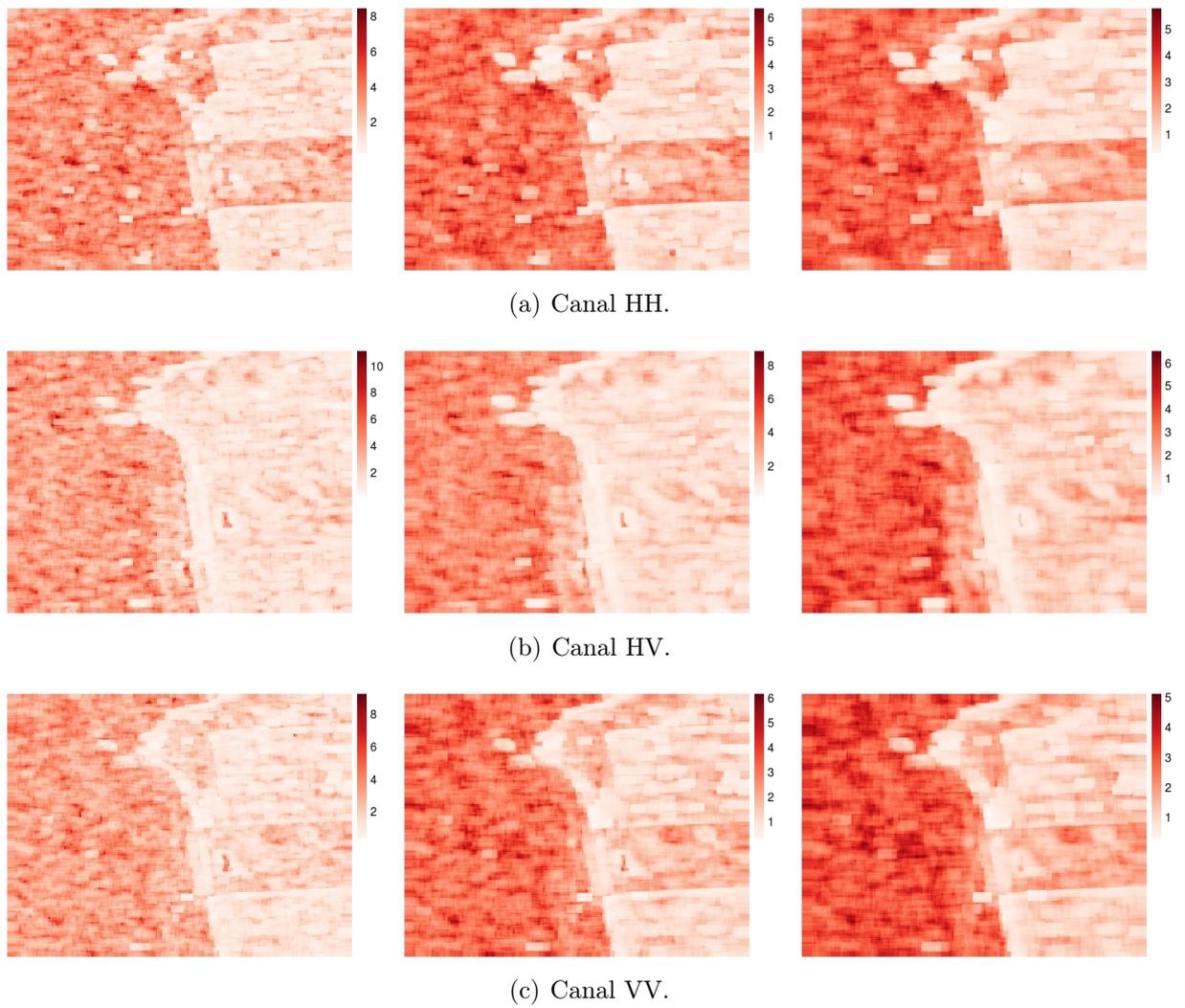


Figura 4.4: Estimativas de L sob o Modelo Γ para $N = 49, 81, 121$, usando janelamentos $7 \times 7, 9 \times 9$ e 11×11 , respectivamente.

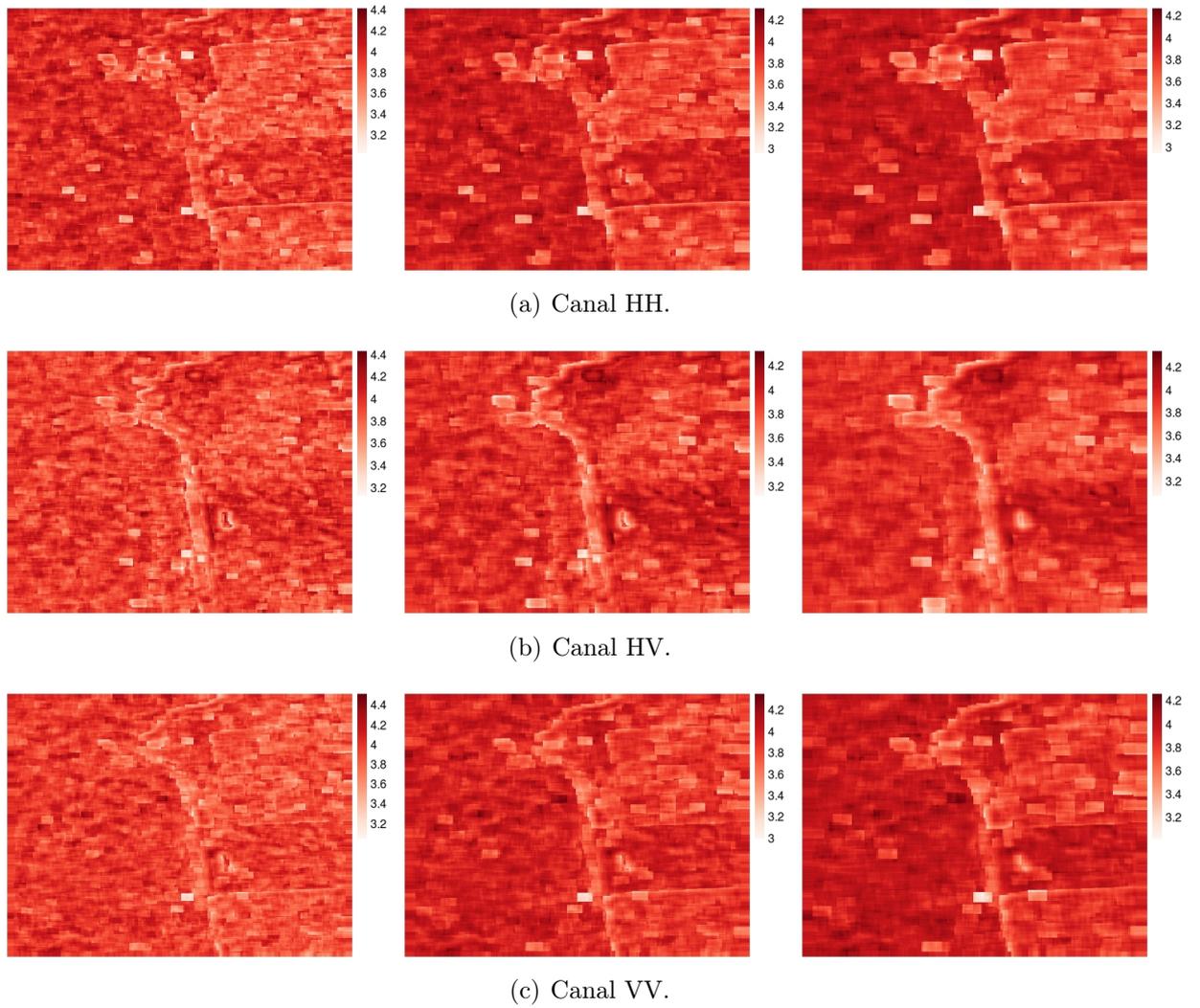


Figura 4.5: Estimativas de L sob o Modelo $G\Gamma$ para $N = 49, 81, 121$, usando janelamentos $7 \times 7, 9 \times 9$ e 11×11 , respectivamente.

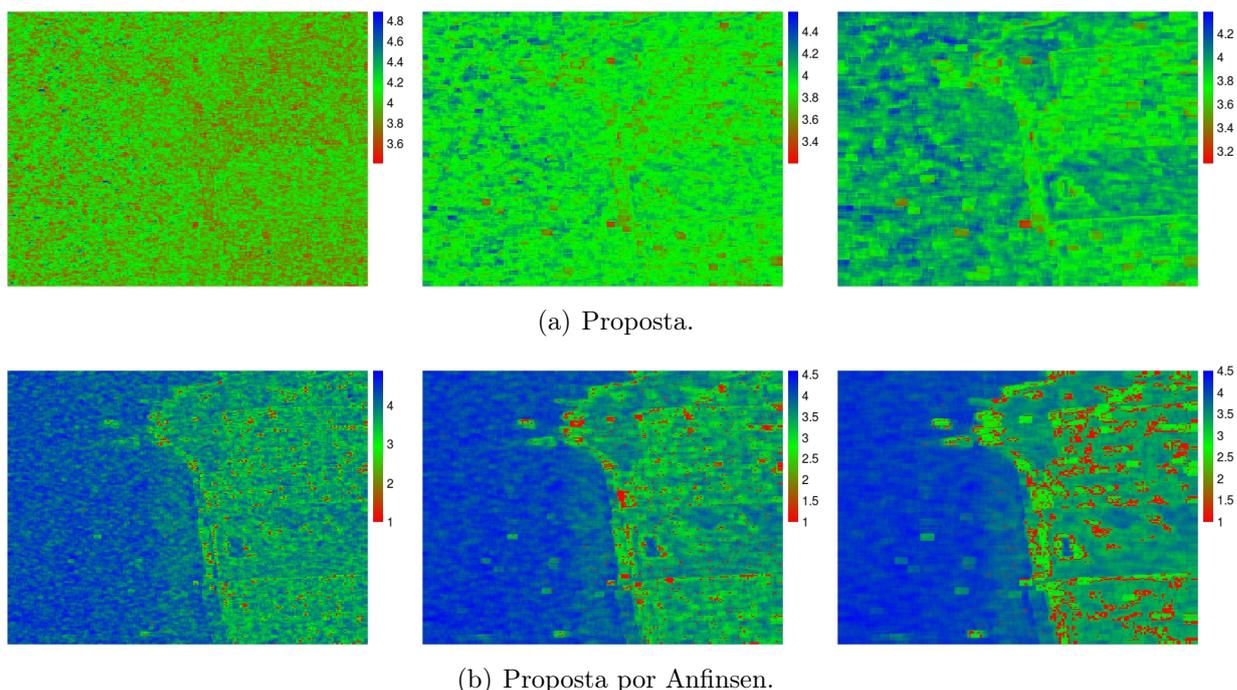


Figura 4.6: Estimativas de L para nossa proposta e proposta por Anfinsen, $N = 9, 25, 49$, usando janelamentos 7×7 , 9×9 e 11×11 , respectivamente.

4.2 Estimação melhorada para o modelo $G\Gamma$: Vieses de segunda ordem de Cox e Snell

4.2.1 Estudo com dados sintéticos

Preliminarmente, um estudo de simulação com dados sintéticos gerados a partir da distribuição $G\Gamma$ foi realizado. Para tanto, 10.000 replicações Monte Carlo são usadas e, para cada, calcula-se:

- (i) Para o estimador MV de θ , $\hat{\theta}$, a mediana dos EMVs e o EQM bem como para os seus respectivos estimadores corrigidos e
- (ii) Para o estimador MV de L , \hat{L} , a mediana dos EMVs e o EQM assim como para os seus respectivos estimadores corrigidos.

Como cenários paramétricos neste estudo de simulação, selecionamos:

- (i) $L \in \{1, 4, 8\}$. O aumento de L implica em imagens resultantes menos afetadas pelo efeito do ruído *speckle*, $L = 1$ reflete o procedimento *singlelook*;
- (ii) $\mu = 1$. Assumimos esta suposição por duas razões: (a) este parâmetro pode ser entendido como uma quantidade conhecida dada pela média amostral das intensidades observadas, e (b) Freitas *et al.* (2005) adotaram $\mu = 1$ para cenários SAR polarimétricos;

- (iii) $\theta > 0$, em que $\theta \in \{0.8, 1, 4\}$. $\theta \geq 1$ representa cenários homogêneos, enquanto $\theta < 1$ caracteriza regiões heterogêneas;

Consideramos o tamanho das amostras $n \in \{49, 144\}$. Também obtivemos os estimadores corrigidos segundo a proposta de Cox-Snell com base na reparametrização da distribuição Γ proposta por Stacy (1962), \hat{L}_{c_1} e $\hat{\theta}_{c_1}$, como para a transformação $Z = L^{-1/\theta}Y$ indicada por Li *et al.* (2011), \hat{L}_{c_2} e $\hat{\theta}_{c_2}$.

A Tabela 4.8 apresenta as estimativas MV para L , \hat{L} , \hat{L}_{c_1} e \hat{L}_{c_2} e seus respectivos EQMs para dados SAR sintéticos. Observando esta tabela, verificamos com base nos EQMs que \hat{L}_{c_1} se torna mais indicado quando n decresce. Para $L = 1$, \hat{L}_{c_2} supera \hat{L}_{c_1} sempre. Considerando $L = 4$, \hat{L}_{c_2} tem melhor desempenho que \hat{L}_{c_1} quando $n = 144$. Por outro lado, para $L = 8$, \hat{L}_{c_1} supera \hat{L}_{c_2} em todos os casos.

Em resumo, \hat{L}_{c_1} funciona bem quando $L \rightarrow \infty$ e $n \rightarrow 0$ em relação a \hat{L}_{c_2} . Assim, \hat{L}_{c_1} pode ser indicado como uma alternativa a \hat{L} para maiores valores de *looks*, quando n decresce.

Tabela 4.8: Estimativas MV para L corrigidas e não corrigidas e seus respectivos EQMs para dados SAR sintéticos

L	θ	n	\hat{L}	EQM	\hat{L}_{c_1}	\hat{L}_{c_2}	EQM $_{\hat{L}_{c_1}}$	EQM $_{\hat{L}_{c_2}}$
1	0.8	49	0.974	0.730	0.976	0.853	0.689	0.244
		144	0.990	0.100	0.991	0.949	0.099	0.080
	1	49	0.978	0.856	0.980	0.856	0.807	0.390
		144	0.993	0.097	0.994	0.952	0.095	0.077
	4	49	0.922	0.628	0.926	0.817	0.592	0.274
		144	0.967	0.093	0.967	0.928	0.091	0.076
4	0.8	49	3.409	9.633	3.344	1.419	9.119	16.584
		144	3.922	6.016	3.895	3.175	5.888	1.660
	1	49	3.310	9.399	3.248	1.408	8.918	15.532
		144	3.863	6.032	3.836	3.139	5.907	1.721
	4	49	3.734	9.877	3.659	1.436	9.289	18.324
		144	4.103	7.008	4.074	3.285	6.847	1.556
8	0.8	49	4.666	22.261	4.563	1.303	22.234	60.643
		144	6.294	12.577	6.241	4.268	12.540	14.628
	1	49	4.381	24.215	4.286	1.355	24.183	59.540
		144	6.215	12.833	6.162	4.242	12.802	14.839
	4	49	5.522	16.866	5.391	1.128	16.821	66.580
		144	6.595	11.778	6.538	4.345	11.714	13.959

A Tabela 4.9 apresenta as estimativas MV para θ , $\hat{\theta}_{c_1}$, $\hat{\theta}_{c_2}$ e seus respectivos EQMs para dados SAR sintéticos. Analisando esta tabela, verificamos com base nos EQMs que $\hat{\theta}_{c_2}$ se torna uma boa alternativa quando n decresce. Para $L = 1$, $\hat{\theta}_{c_1}$ supera $\hat{\theta}_{c_2}$. Considerando $L \in \{4, 8\}$, $\hat{\theta}_{c_2}$ tem melhor desempenho que $\hat{\theta}_{c_1}$ em todos os cenários.

Enfim, $\hat{\theta}_{c_2}$ funciona bem quando $L \rightarrow \infty$ e $n \rightarrow 0$ em relação a $\hat{\theta}_{c_1}$. Logo, $\hat{\theta}_{c_2}$ pode ser indicado como uma alternativa a $\hat{\theta}$, sendo as diferenças mais significativas em pequenas amostras.

Como mencionado anteriormente, os vieses são substanciais em métodos que requerem menores tamanhos amostrais. Neste sentido, os estimadores corrigidos, diga-se $\hat{\theta}_{c_2}$ e \hat{L}_{c_1} , destacados neste estudo simulado possuem melhor desempenho que os estimadores não corrigidos. Ressalta-se que os EMVs corrigidos e não corrigidos são assintoticamente semelhantes. Portanto, a proposta de corrigir o viés fornece estimadores melhorados.

Tabela 4.9: Estimativas MV para θ corrigidas e não corrigidas e seus respectivos EQMs para dados SAR sintéticos

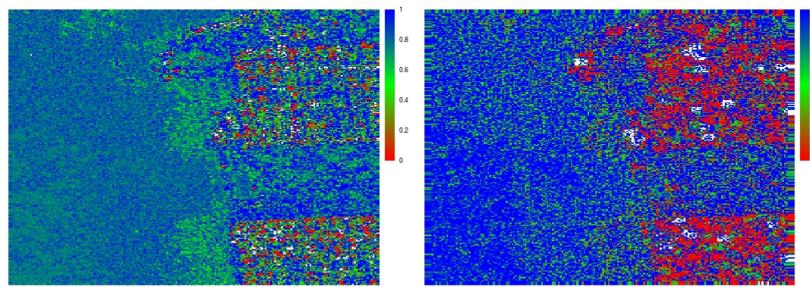
L	θ	n	$\hat{\theta}$	EQM	$\hat{\theta}_{c_1}$	$\hat{\theta}_{c_2}$	EQM $_{\hat{\theta}_{c_1}}$	EQM $_{\hat{\theta}_{c_2}}$
	0.8	49	0.818	0.073	0.789	0.770	0.057	0.062
		144	0.805	0.018	0.796	0.789	0.017	0.017
1	1	49	1.023	0.115	0.987	0.963	0.089	0.098
		144	1.004	0.028	0.992	0.984	0.026	0.027
	4	49	4.180	1.411	4.019	3.930	1.106	1.166
		144	4.072	0.472	4.023	3.990	0.438	0.441
	0.8	49	0.878	0.144	0.865	0.839	0.131	0.122
		144	0.811	0.045	0.807	0.798	0.044	0.043
4	1	49	1.111	0.240	1.095	1.062	0.218	0.203
		144	1.019	0.071	1.014	1.004	0.070	0.068
	4	49	4.282	2.141	4.219	4.091	1.987	1.833
		144	4.002	1.145	3.983	3.943	1.124	1.104
	0.8	49	1.066	0.288	1.051	1.020	0.265	0.240
		144	0.905	0.070	0.901	0.892	0.069	0.066
8	1	49	1.357	0.462	1.338	1.298	0.426	0.384
		144	1.137	0.114	1.132	1.121	0.111	0.106
	4	49	4.872	2.610	4.812	4.669	2.428	2.108
		144	4.439	1.419	4.421	4.377	1.386	1.326

4.2.2 Aplicação com dados SAR de intensidade

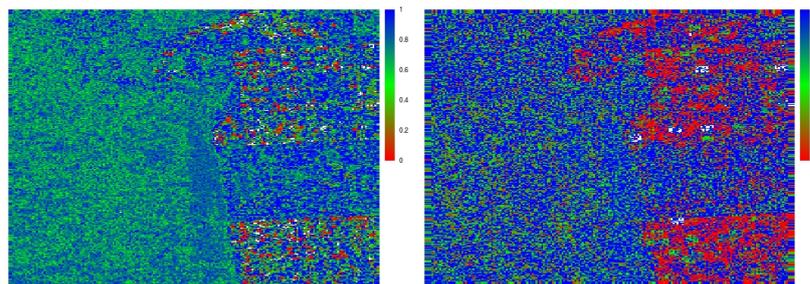
Nesta subseção, complementamos o estudo de simulação com um experimento para dados SAR de intensidade, objetivando destacar o uso da correção de viés como método aperfeiçoado de estimação MV para o modelo GF . Analogamente a Seção 4.1, consideramos as mesmas três regiões da imagem da Baía de São Francisco observada na Figura 4.2(a), obtida pelo sensor AIRSAR na banda-L com número nominal de *looks* $L = 4$.

Nesta etapa, o TRV é razoável para testar $\mathcal{H}_0: \text{GF} = \text{GF}_{\text{cor}}$, ou, equivalentemente, $\hat{\theta} = \hat{\theta}_c$. Particularmente, sob \mathcal{H}_0 , a estatística $S_{\text{RV}} \sim \chi_3^2$, em que $S_{\text{RV}} = -2\{\ell(\hat{\theta}_c) - \ell(\hat{\theta})\}$, pode ser entendida como uma medida de distância entre os modelos GF e GF corrigido. A Figura 4.7 apresenta mapas de p -valor para a estatística S_{RV} nos canais HH e VV. Nota-se que os p -valores diminuem quando as texturas se intensificam, em outras palavras, as diferenças entre as log-verossimilhanças GF e GF corrigida aumentam. Assim, os dados fornecem evidência da existência de diferenças (distâncias) significativas entre os modelos corrigido e não corrigido principalmente em cenários urbanos. A mesma conclusão é obtida ao analisar a Figura 4.7 quando $n \rightarrow \infty$. Em contrapartida, em regiões oceânicas, a

diferença não é expressiva.



(a) Canal HH.



(b) Canal VV.

Figura 4.7: p -valores para os testes da razão entre verossimilhanças sob $\mathcal{H}_0: G\Gamma = G\Gamma_{cor}$ e $N = 9, 49$, para janelamentos 3×3 e 7×7 , respectivamente.

4.3 Medidas de divergência com base nas distribuições $G\Gamma$ e WCE

4.3.1 Estudo sintético

Um estudo de simulação foi realizado envolvendo dados gerados a partir das distribuições $G\Gamma$ e Γ . Os dados sintéticos foram utilizados como entrada para os cinco testes de hipóteses de duas amostras propostos na Proposição 1: S_{DZY} , S_{DYZ} , S_{dYZ} , S_{dZZ} e S_{dYY} . Para tal, considera-se 10.000 replicações Monte Carlo para quantificar as seguintes figuras a fim de mérito:

- tamanho e poder dos testes ao nível de significância $\eta \in \{1\%, 5\%\}$ e
- média das estatísticas do teste.

A configuração paramétrica adotada é dada por:

- $L \in \{1, 3, 4, 8\}$ (quando L aumenta, a imagem resultante é menos afetada pelo efeito do ruído *speckle*);
- $\mu = 1$. Assumimos esta suposição por duas razões: (a) a variação dos valores de μ não afetam as estimativas do erro de tipo I e (b) Freitas *et al.* (2005) adotaram $\mu = 1$ para cenários SAR polarimétricos.

(iii) $|\theta| = 1$, $|\theta| < 1$ e $|\theta| > 1$, em que $|\theta| \in \{0.1, 0.5, 0.8, 1, 2, 5, 8\}$.

Figura 4.8 exibe imagens sintéticas para alguns pontos paramétricos do modelo Γ . Nota-se que, cenários pouco rugosos estão associados com valores menores de θ e maiores valores de L . Este resultado para θ é esperado porque Γ implica $\Gamma\Gamma$ a partir da transformação $1/\theta$, que tende à distribuição degenerada se $|\theta| \rightarrow \infty$ e gera uma variável com valores maiores quando $|\theta| \rightarrow 0$. Adotamos também uma variação no tamanho das amostras $n \in \{64, 81, 100, 121, 144\}$.

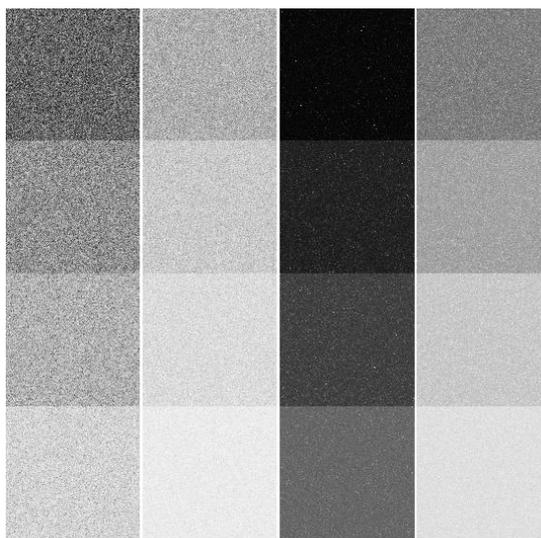


Figura 4.8: Imagem sintética $\Gamma\Gamma$ para $(L, \text{sign}(\theta)) \in \{(1, 1), (8, 1), (1, -1), (8, -1)\}$ e $\theta \in \{3, 5, 8, 15\}$.

Em resumo, seguimos os passos dados no Algoritmo 2.

Algorithm 2 Esquema para o estudo empírico

- 1: Para cada replicação Monte Carlo, $k = 1, 2, \dots, MC$, faça:
- 2: Gere duas amostras $\Gamma\Gamma$ distribuídas e—para ambas—calcule estimativas MV $\hat{\delta}_{k,i} = (\hat{\theta}_{k,i}, \hat{\mathcal{L}}_{k,i}, \hat{\mu}_{k,i})^\top$ for $i = 1, 2$.
- 2: Calcule as medidas de contraste discutidas, $S_v^{(k)}(\hat{\delta}_{k,1}, \hat{\delta}_{k,2})$, para $v \in \{DZY, DYZ, dYZ, dZZ, dYY\}$, bem como as respectivas estatísticas de teste para cada replicação de Monte Carlo.
- 3: Defina a variável binária $\mathbb{I}_{k,v} = \mathbb{I}_{(\chi_{1-\eta,3}^2, \infty)}(S_v^{(k)})$.
- 4: Calcule

$$\hat{\alpha}_v = \frac{1}{MC} \sum_{k=1}^{MC} \mathbb{I}_{k,v} \text{ and } \bar{S}_v = \frac{1}{MC} \sum_{k=1}^{MC} S_v^{(k)}.$$

As Tabelas 4.10-4.12 apresentam os tamanhos de teste empíricos em níveis nominais de 1% e 5% e a média das estimativas para as estatísticas de Kullback-Leibler ($d_{\text{KL}}(Y_1, Y_2)$ e $d_{\text{KL}}(Z_1, Z_2)$). Embora tenha ocorrido leves oscilações para pequenos tamanhos de amostra, taxas de rejeição sob a hipótese nula para θ e N fixos diminuem à medida que L aumenta, ou, equivalentemente, quando o efeito do ruído na imagem sintética diminui. Em termos

estatísticos, a média das estatísticas tende a assumir o valor de três (grau de liberdade da distribuição assintótica das estatísticas) quando se aumenta ambos o tamanho da amostra e/ou o número de *looks*. Isso pode ser explicado pela *lei dos grandes números*. Em geral, os resultados indicam que as medidas de contraste estocástico entre distribuições $G\Gamma$ são melhores do que aquela em termos dos modelos Γ . Pode-se notar ainda que as medidas de contraste para $\theta \rightarrow 0^-$ apresentaram desempenho ruim, mas contrastes baseados em $G\Gamma$ superam aqueles com base em Γ . Em resumo, o contraste $G\Gamma$ é mais indicado.

Tabela 4.10: Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ em que $|\theta| = 1, \mu = 1, L \in \{1, 3, 4, 8\}$ e tamanho amostral N

(θ, μ)	N	nível nominal 1%				nível nominal 5%				\bar{S}_\bullet			
		L				L				1	3	4	8
<i>Para a distribuição Γ</i>													
(1, 1)	64	0.0512	0.0445	0.0435	0.0470	0.1160	0.1038	0.1014	0.1061	4.0481	3.7739	3.7631	3.7457
	81	0.0430	0.0348	0.0367	0.0341	0.1044	0.0901	0.0946	0.0861	3.7714	3.5564	3.5821	3.5099
	121	0.0291	0.0230	0.0224	0.0228	0.0878	0.0767	0.0740	0.0675	3.5239	3.3172	3.2808	3.2072
	144	0.0234	0.0196	0.0176	0.0153	0.0805	0.0654	0.0678	0.0612	3.4119	3.2225	3.2035	3.0886
	225	0.0194	0.0166	0.0154	0.0123	0.0664	0.0651	0.0633	0.0523	3.2533	3.1617	3.1587	2.9939
<i>Para a distribuição recíproca Γ</i>													
(-1, 1)	64	0.0484	0.0445	0.0433	0.0448	0.1111	0.1040	0.1012	0.1064	3.9667	3.7767	3.7585	3.7696
	81	0.0435	0.0348	0.0331	0.0363	0.1038	0.0902	0.0867	0.0924	3.7637	3.5569	3.5456	3.5667
	121	0.0290	0.0230	0.0202	0.0211	0.0877	0.0767	0.0729	0.0718	3.5230	3.3172	3.2697	3.2021
	144	0.0234	0.0196	0.0177	0.0173	0.0805	0.0654	0.0662	0.0569	3.4123	3.2225	3.1640	3.0724
	225	0.0194	0.0166	0.0156	0.0139	0.0664	0.0651	0.0605	0.0498	3.2533	3.1617	3.1279	3.0203

Adicionalmente, realiza-se um breve estudo do poder do teste para as medidas propostas. Os resultados são apresentados na Figura 4.9. Nota-se que as medidas propostas podem ser muito sensíveis para detectar mudança do ponto de vista do contraste (que pode ser usado como solução para problemas de processamento de imagem, tais como em segmentação e classificação). Com relação à comparação entre os modelos Γ e $G\Gamma$, pode-se observar:

- Para $\theta < 1$: o erro ao escolher o modelo Γ quando a verdadeira distribuição (em processamento de imagem SAR, valor de referência) é $G\Gamma$, $D_{KL}(Z||Y)$, é maior que sua contrapartida $D_{KL}(Y||Z)$. Em outras palavras, $G\Gamma$ é mais recomendado para dados com alta variabilidade.
- Para $\theta > 1$: Analogamente, $D_{KL}(Z||Y) < D_{KL}(Y||Z)$ aponta que o modelo Γ é mais indicado para dados homogêneos.

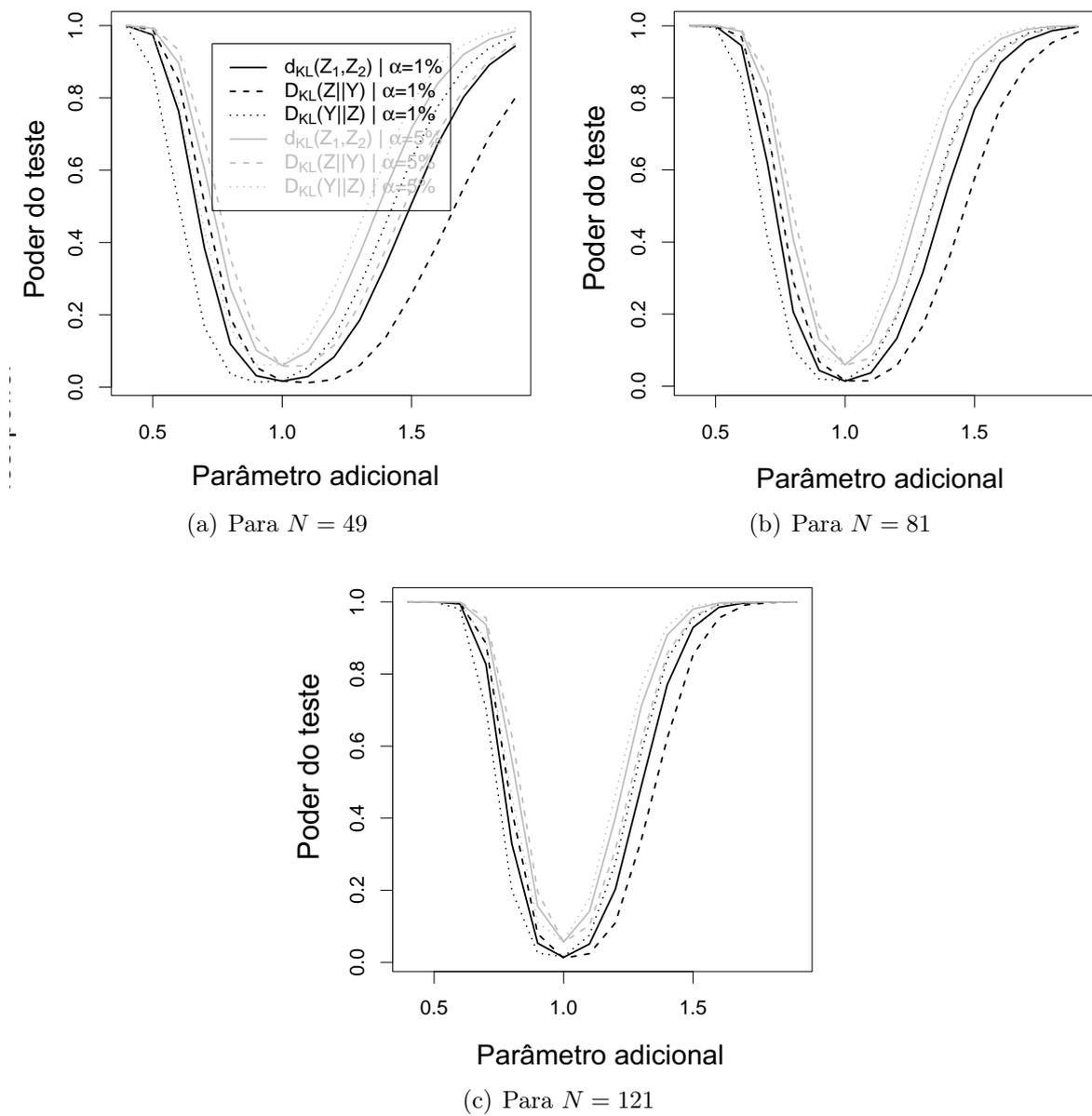


Figura 4.9: Curvas de poder do teste para a estatística proposta.

4.3.2 Aplicação com dados SAR de intensidade

Nesta subseção, realizamos uma aplicação para a imagem de São Francisco, obtida pelo sensor AIRSAR com quatro números de equivalentes em *looks*, para tal considera-se novamente a Figura 4.2(a).

Figura 4.10 apresenta os valores estimados de θ na imagem AIRSAR considerada. Em geral, pode-se observar que estimativas menores estão associados com cenários mais texturados. Este comportamento foi mencionado na discussão da imagem sintética na Figura 4.8.

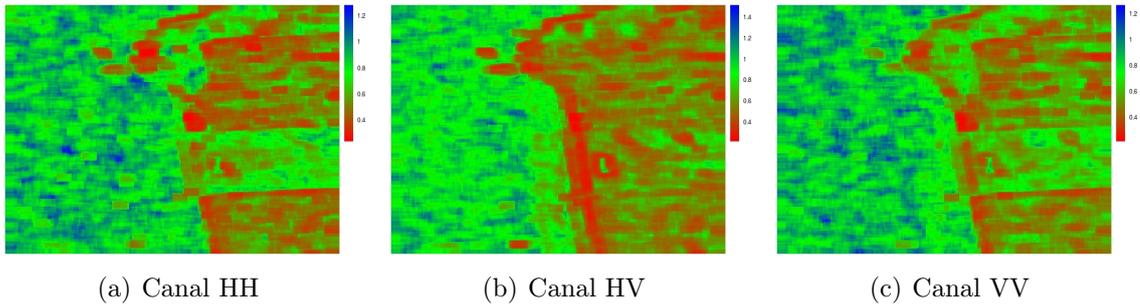


Figura 4.10: Mapa das estimativas de θ para ajustes Γ e $G\Gamma$ em janela 10×10 .

Adaptando o Algoritmo 2 a dados reais, quantificou-se o desempenho dos métodos. As Tabelas 4.13-4.15 exibem as taxas de rejeição (taxas de alarme falso) dos testes de hipóteses baseados nas estatísticas S_{dZZ} (distâncias entre variáveis $G\Gamma$) e S_{dYY} (distâncias entre variáveis Γ). O teste em função de S_{dZZ} supera aquele S_{dYY} em todos os casos, indicando que o uso das discrepâncias entre modelos $G\Gamma$ é mais informativo que aqueles entre distribuições Γ . Em particular, a discrepância S_{dYY} funcionou bem nas regiões oceânicas nos canais VV e HV.

Com exceção da região HH de florestas e das áreas urbanas, as taxas tendem a melhorar rapidamente aumentando o tamanho da janela. Por outro lado, Tabela 4.13 mostra que o contraste entre modelos $G\Gamma$ é significativamente melhor do que entre Γ para as áreas de floresta e zona urbana sendo assintoticamente similar à discrepância entre Γ em regiões oceânicas. Em resumo, o contraste entre pixels Γ distribuídos é recomendado apenas para regiões de oceano com pequenos tamanhos de amostra.

Nota-se que o teste de hipóteses Υ_{dZZ} baseado nas estatísticas S_{dZZ} avalia a hipótese $\mathcal{H}_0: [\theta_1, L_1, \mu_1] = [\theta_2, L_2, \mu_2] \times \mathcal{H}_1: [\theta_1, L_1, \mu_1] \neq [\theta_2, L_2, \mu_2]$. Sua região crítica é dada por

$$CR = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n+m} : S_{dZZ}(\mathbf{x}, \mathbf{y}) > \chi_{\alpha,3}^2\}$$

e sua função poder-dada $\boldsymbol{\delta} = [\theta_1, L_1, \mu_1, \theta_2, L_2, \mu_2]^\top \in \Theta \subseteq \mathbb{R}^6$ é definida por

$$\pi(\boldsymbol{\delta}) = \Pr((\mathbf{x}, \mathbf{y}) \in CR \mid \boldsymbol{\delta} \in \Theta).$$

O segundo objetivo desta subseção é comparar as distribuições Γ (modelagem padrão

para descrever o ruído *speckle*) e Γ . As suposições $\mu_1 = \mu_2 = 1$ (FREITAS *et al.*, 2005) e $L_1 = L_2 = 4$ (como resultado desta aplicação) são razoáveis na prática. Assim, a fim de quantificar $\pi(\boldsymbol{\delta})$, definimos $\theta_1 = 1$ e $\theta_2 = \theta > 0$.

Também empregamos o contraste S_{dYZ} , S_{dZY} e S_{dYZ} na vizinhança de cada pixel da Figura 1.3. Os resultados são apresentados na Figura 4.11. Pode-se notar que essas medidas indicam mudanças abruptas de texturas. Adicionalmente, aplicamos o contraste

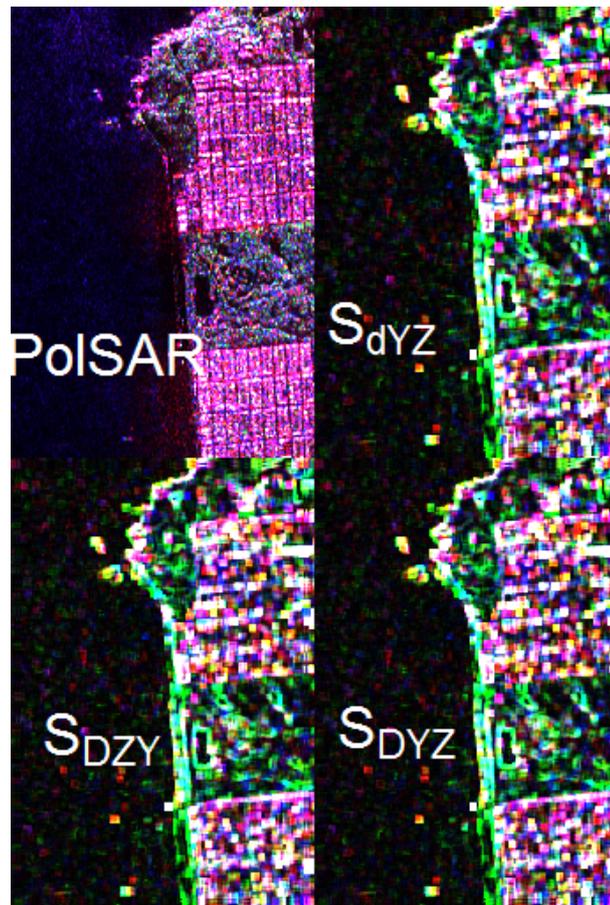
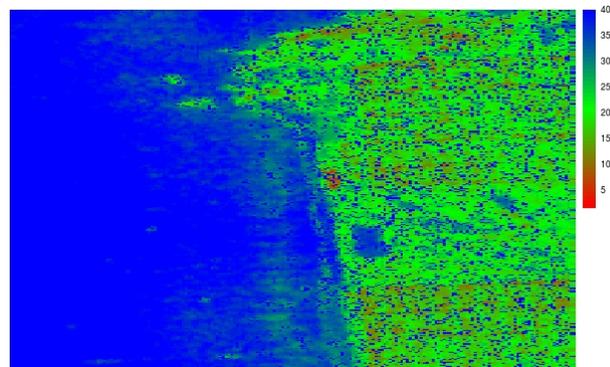
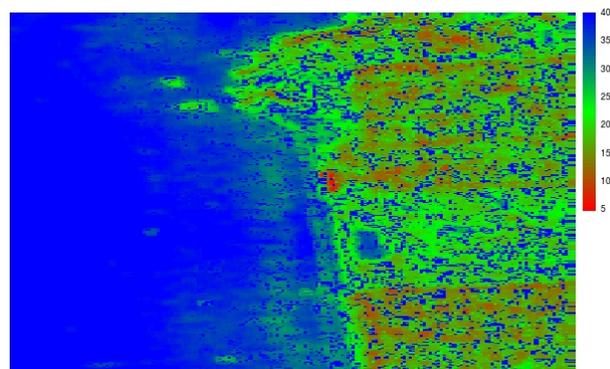
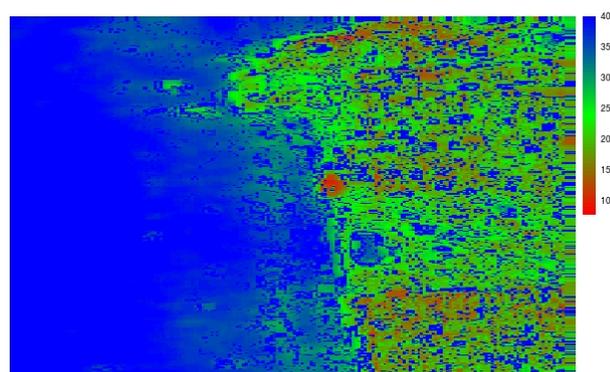
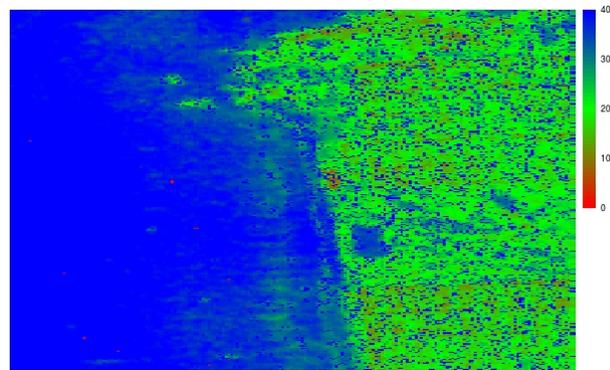
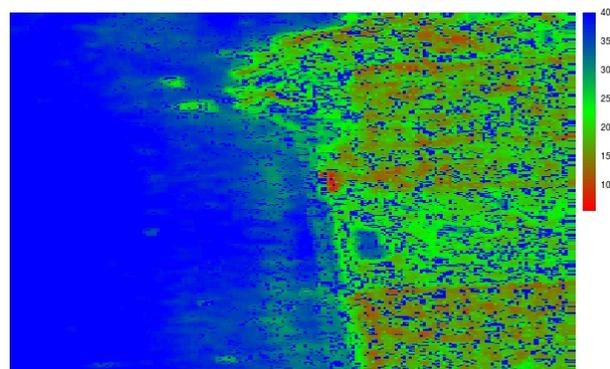
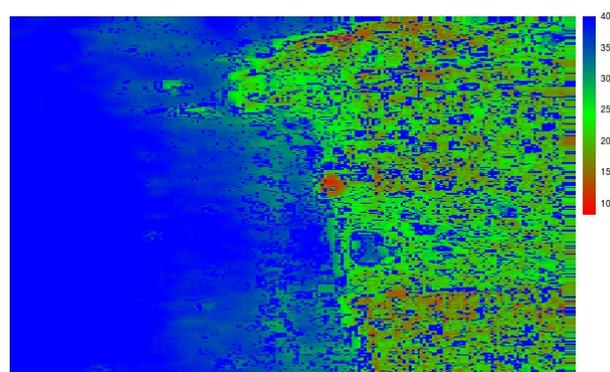


Figura 4.11: Mapa dos valores de S_{DZY} , S_{DYZ} e S_{dZZ} com base em uma janela 10×10 .

S_{ZZ} em uma imagem real PolSAR para janelas 10×10 ao redor de seus pixels. Resultados são apresentados nas Figuras 4.12-4.15. Como esperado, os valores são maiores do que os aguardados dos quantis χ^2 . Contudo, para alguns tamanhos de amostra, o contraste entre modelos WCE e Γ indica que a hipótese de correlação nula entre os canais pode ser encontrada em cenários com textura pronunciada.

(a) Para $N = 9$ (b) Para $N = 25$ (c) Para $N = 49$ Figura 4.12: Divergência KL entre os modelos Γ e \mathcal{W}

(a) Para $N = 9$ (b) Para $N = 25$ (c) Para $N = 49$ Figura 4.13: Divergência KL entre os modelos Γ e \mathcal{W}

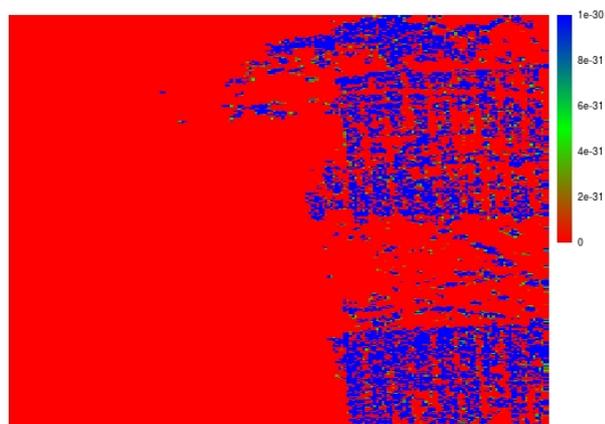
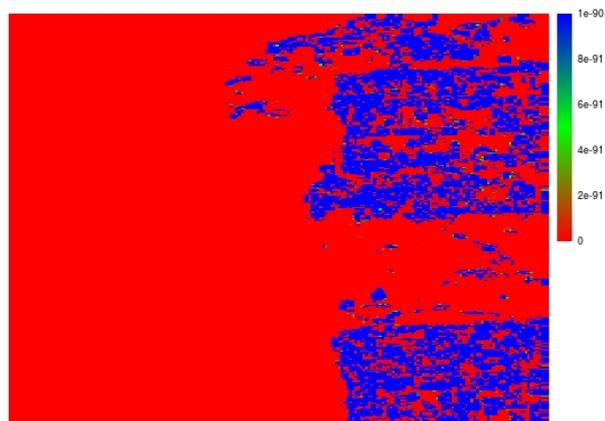
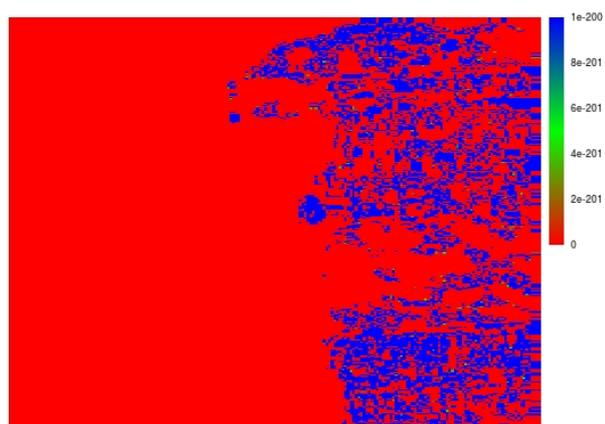
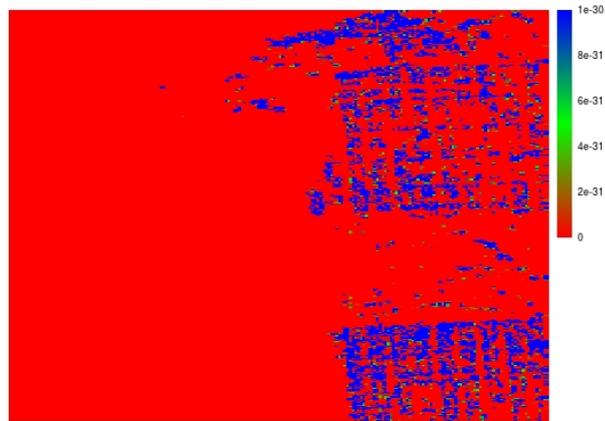
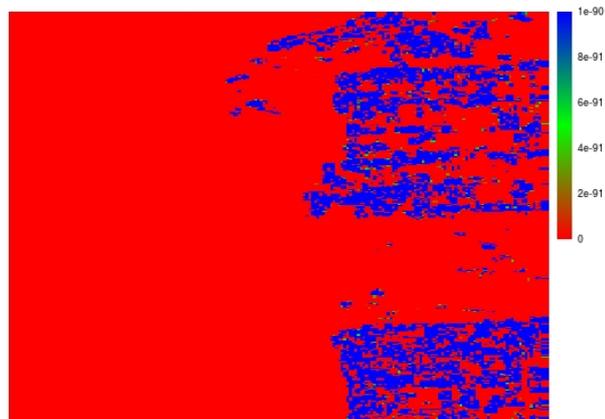
(a) Para $N = 9$ (b) Para $N = 25$ (c) Para $N = 49$

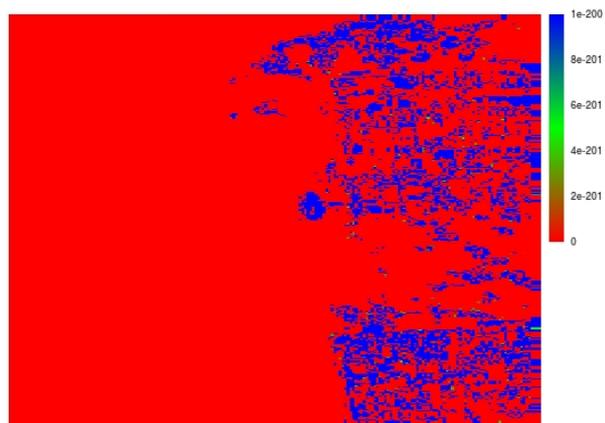
Figura 4.14: p -valor do teste de hipótese baseado na divergência KL entre os modelos $G\Gamma$ e \mathcal{W}



(a) Para $N = 9$



(b) Para $N = 25$



(c) Para $N = 49$

Figura 4.15: p -valor do teste de hipótese baseado na divergência KL entre os modelos Γ e \mathcal{W}

Tabela 4.11: Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ em que $|\theta| < 1, \mu = 1, L \in \{1, 3, 4, 8\}$ e tamanho amostral N

(θ, μ)	N	nível nominal 1%				nível nominal 5%				\bar{S}_\bullet			
		L				L				1	3	4	8
<i>Para a distribuição GF sob $\theta \in (0, 1)$</i>													
(0.1, 1)	64	0.0524	0.0284	0.0233	0.0256	0.1202	0.0808	0.0715	0.0683	4.0540	3.4135	3.3094	3.1699
	81	0.0434	0.0290	0.0221	0.0251	0.1006	0.0798	0.0722	0.0698	3.8223	3.3938	3.2406	3.1887
	121	0.0293	0.0185	0.0195	0.0231	0.0850	0.0684	0.0653	0.0675	3.5161	3.2626	3.1665	3.1663
	144	0.0241	0.0187	0.0201	0.0145	0.0804	0.0661	0.0663	0.0613	3.3978	3.2181	3.2083	3.0827
	225	0.0189	0.0159	0.0151	0.0115	0.0749	0.0639	0.0612	0.0509	3.3135	3.1787	3.1551	2.9554
(0.5, ●)	64	0.0536	0.0436	0.0429	0.0414	0.1216	0.1046	0.0995	0.1029	4.1049	3.8064	3.7008	3.7243
	81	0.0449	0.0362	0.0342	0.0382	0.1063	0.0909	0.0932	0.0944	3.8726	3.5667	3.5203	3.5735
	121	0.0306	0.0200	0.0208	0.0223	0.0854	0.0731	0.0703	0.0709	3.4940	3.2800	3.2420	3.1869
	144	0.0269	0.0194	0.0169	0.0185	0.0849	0.0671	0.0692	0.0669	3.4368	3.2412	3.1970	3.1282
	225	0.0168	0.0163	0.0165	0.0131	0.0705	0.0616	0.0633	0.0557	3.2422	3.1729	3.1438	3.0389
(0.8, ●)	64	0.0531	0.0433	0.0420	0.0445	0.1196	0.1037	0.1056	0.1056	4.0863	3.7202	3.7507	3.8003
	81	0.0479	0.0332	0.0345	0.0341	0.1110	0.0869	0.0905	0.0921	3.8718	3.5706	3.5351	3.5623
	121	0.0287	0.0232	0.0220	0.0238	0.0890	0.0760	0.0711	0.0728	3.5261	3.3374	3.2356	3.1865
	144	0.0255	0.0188	0.0198	0.0181	0.0805	0.0686	0.0664	0.0655	3.3969	3.2453	3.2271	3.0973
	225	0.0193	0.0155	0.0159	0.0128	0.0734	0.0642	0.0635	0.0504	3.3319	3.1520	3.1906	2.9774
<i>Para a distribuição GF sob $\theta \in (-1, 0)$</i>													
(-0.1, 1)	64	0.2775	0.2885	0.2659	0.2609	0.3557	0.3497	0.3330	0.3265	9.4063	9.5091	9.0461	8.8521
	81	0.2900	0.2838	0.2735	0.2523	0.3465	0.3575	0.3429	0.3375	9.4180	9.3840	9.0986	8.8710
	121	0.2551	0.2836	0.2555	0.0292	0.3150	0.3619	0.3370	0.0710	8.5863	9.4803	8.7296	3.2773
	144	0.2821	0.3046	0.2089	0.0177	0.3611	0.3580	0.2649	0.0614	8.8669	9.7323	7.5114	3.0692
	225	0.2612	0.0259	0.0169	0.0123	0.3483	0.0637	0.0634	0.0533	9.0691	3.2469	3.1614	2.9931
(-0.5, ●)	64	0.2381	0.0430	0.0412	0.0452	0.3144	0.0994	0.1003	0.1049	8.4738	3.6888	3.6882	3.7406
	81	0.1123	0.0335	0.0345	0.0328	0.1703	0.0878	0.0888	0.0907	5.5570	3.5354	3.5272	3.5346
	121	0.0283	0.0211	0.0219	0.0241	0.0837	0.0693	0.0733	0.0708	3.4528	3.2535	3.2910	3.2188
	144	0.0230	0.0194	0.0198	0.0188	0.0755	0.0659	0.0682	0.0661	3.3505	3.2112	3.2206	3.1254
	225	0.0192	0.0151	0.0154	0.0128	0.0719	0.0631	0.0620	0.0505	3.2593	3.1629	3.1424	2.9691
(-0.8, ●)	64	0.0509	0.0409	0.0462	0.0459	0.1166	0.0995	0.1057	0.1113	4.0281	3.6803	3.7282	3.8102
	81	0.0461	0.0352	0.0365	0.0358	0.1103	0.0892	0.09431	0.0931	3.8691	3.5753	3.5832	3.5393
	121	0.0283	0.0240	0.0194	0.0207	0.0862	0.0742	0.0707	0.0722	3.5202	3.3063	3.2424	3.1833
	144	0.0255	0.0179	0.0177	0.0165	0.0802	0.0668	0.0679	0.0635	3.3928	3.2311	3.2221	3.1177
	225	0.0196	0.0159	0.0152	0.0134	0.0732	0.0624	0.0595	0.0548	3.2842	3.1681	3.1506	3.0069

Tabela 4.12: Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ em que $|\theta| > 1, \mu = 1, L \in \{1, 3, 4, 8\}$ e tamanho amostral N

(θ, μ)	N	nível nominal 1%				nível nominal 5%				\bar{S}_\bullet			
		L				L				1	3	4	8
<i>Para a distribuição GF sob $\theta \in (1, \infty)$</i>													
(2, 1)	64	0.0544	0.0401	0.0401	0.0465	0.1257	0.0984	0.1008	0.1086	4.1163	3.6629	3.6811	3.8291
	81	0.0383	0.0346	0.0372	0.0340	0.1000	0.0870	0.0963	0.0939	3.7347	3.5578	3.5682	3.5275
	121	0.0325	0.0208	0.0224	0.0208	0.0892	0.0702	0.0703	0.0698	3.5166	3.3115	3.2485	3.1759
	144	0.0226	0.0184	0.0179	0.0169	0.0759	0.0682	0.0640	0.0641	3.3741	3.2441	3.1900	3.1113
	225	0.0203	0.016	0.0152	0.0122	0.0708	0.0617	0.0633	0.0518	3.3239	3.1754	3.1572	2.9944
(5, •)	64	0.0459	0.0366	0.0368	0.0337	0.1102	0.0950	0.0955	0.0850	3.9306	3.6068	3.5552	3.4852
	81	0.0429	0.0288	0.0275	0.0306	0.1034	0.0830	0.0836	0.0844	3.7650	3.4551	3.4015	3.4188
	121	0.0292	0.0225	0.0216	0.0211	0.0869	0.0734	0.0705	0.0675	3.4587	3.3013	3.2532	3.1529
	144	0.0238	0.0224	0.0192	0.0178	0.0790	0.0711	0.0702	0.0596	3.4249	3.2588	3.2130	3.0605
	225	0.0197	0.0168	0.0165	0.0128	0.0694	0.0648	0.0632	0.0543	3.2719	3.1799	3.1712	3.0135
(8, •)	64	0.0281	0.0225	0.0235	0.0191	0.0864	0.0637	0.0672	0.0596	3.4748	3.1099	3.1274	2.9325
	81	0.0249	0.0187	0.0186	0.0185	0.0812	0.0612	0.0576	0.0588	3.4118	3.1295	3.0468	2.8809
	121	0.0235	0.0157	0.0148	0.0112	0.0786	0.0616	0.0577	0.0457	3.3608	3.1729	3.0663	2.7469
	144	0.0223	0.016	0.0124	0.0105	0.0795	0.0617	0.0553	0.0422	3.3337	3.1412	3.0336	2.7227
	225	0.0185	0.0151	0.0118	0.0088	0.0718	0.0605	0.0561	0.0443	3.2982	3.1419	3.0768	2.8504
<i>Para a distribuição GF sob $\theta \in (-\infty, -1)$</i>													
(-2, 1)	64	0.0543	0.0401	0.0401	0.0465	0.1256	0.0984	0.1008	0.1086	4.1133	3.6629	3.6811	3.8291
	81	0.0383	0.0346	0.0372	0.0339	0.1000	0.0869	0.0963	0.0939	3.7352	3.5578	3.5682	3.5275
	121	0.0325	0.0208	0.0224	0.0209	0.0892	0.0702	0.0703	0.0695	3.5166	3.3115	3.2485	3.1791
	144	0.0226	0.0184	0.0179	0.0174	0.0759	0.0682	0.0640	0.0654	3.3741	3.2441	3.1900	3.1200
	225	0.0203	0.016	0.0152	0.0119	0.0708	0.0617	0.0633	0.0526	3.3238	3.1754	3.1572	3.0244
(-5, •)	64	0.0459	0.0366	0.0367	0.0335	0.1102	0.0951	0.0955	0.0883	3.9306	3.6068	3.5551	3.4900
	81	0.0429	0.0288	0.0275	0.0287	0.1034	0.0830	0.0837	0.0836	3.7650	3.4551	3.4015	3.3448
	121	0.0292	0.0225	0.0216	0.0220	0.0869	0.0734	0.0705	0.0679	3.4587	3.3013	3.2532	3.1684
	144	0.0238	0.0224	0.0192	0.0170	0.0790	0.0711	0.0702	0.0591	3.4249	3.2588	3.2130	3.0833
	225	0.0197	0.0168	0.0165	0.0090	0.0694	0.0648	0.0632	0.0483	3.2719	3.1799	3.1711	2.9239
(-8, •)	64	0.0281	0.0226	0.0233	0.0212	0.0864	0.0638	0.0686	0.0611	3.4748	3.1122	3.1331	2.9855
	81	0.0249	0.0193	0.0189	0.0187	0.0812	0.0596	0.0588	0.0603	3.4118	3.0837	3.0331	2.9045
	121	0.0235	0.0174	0.0135	0.0129	0.0786	0.0636	0.0539	0.0520	3.3608	3.1644	3.0238	2.8326
	144	0.0223	0.0157	0.0150	0.0128	0.0795	0.0603	0.0559	0.0463	3.3337	3.1505	3.0606	2.7428
	225	0.0185	0.0151	0.0137	0.0085	0.0718	0.0602	0.0592	0.0442	3.2982	3.1317	3.0861	2.8066

Tabela 4.13: Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ e tamanho amostral N para distribuições GF baseada nos dados

Canal	Região	nível nominal 1%					nível nominal 5%					\bar{S}_\bullet							
		64	81	100	121	144	225	64	81	100	121	144	225	64	81	100	121	144	225
HH	Oceano	0.071	0.058	0.050	0.040	0.040	0.024	0.150	0.132	0.117	0.108	0.099	0.078	4.378	4.098	3.899	3.738	3.657	3.289
	Floresta	0.145	0.135	0.125	0.124	0.120	0.100	0.247	0.241	0.244	0.240	0.240	0.220	5.856	5.710	5.604	5.522	5.390	4.890
	Urbana	0.042	0.023	0.012	0.008	0.006	0.002	0.091	0.062	0.043	0.035	0.028	0.011	3.381	3.028	2.759	2.579	2.429	2.000
HV	Oceano	0.040	0.025	0.018	0.011	0.009	0.004	0.091	0.070	0.060	0.043	0.040	0.027	3.467	3.179	2.948	2.799	2.703	2.505
	Floresta	0.027	0.015	0.010	0.006	0.003	0.001	0.074	0.053	0.037	0.026	0.021	0.006	3.138	2.848	2.559	2.419	2.271	1.848
	Urbana	0.070	0.050	0.033	0.026	0.018	0.006	0.139	0.109	0.082	0.071	0.054	0.028	4.173	3.780	3.463	3.242	2.976	2.449
VV	Oceano	0.046	0.032	0.026	0.020	0.017	0.009	0.103	0.084	0.070	0.062	0.052	0.037	3.679	3.407	3.189	3.023	2.947	2.688
	Floresta	0.053	0.037	0.030	0.025	0.018	0.005	0.119	0.102	0.089	0.080	0.066	0.030	3.879	3.663	3.442	3.273	3.039	2.439
	Urbana	0.056	0.029	0.017	0.010	0.007	0.002	0.110	0.070	0.056	0.038	0.028	0.011	3.525	3.139	2.885	2.624	2.476	2.009

Tabela 4.14: Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ e tamanho amostral N para distribuições $GF|L = 4$ baseada nos dados

Canal	Região	nível nominal 1%						nível nominal 5%						\bar{S}_\bullet					
		64	81	100	121	144	225	64	81	100	121	144	225	64	81	100	121	144	225
HH	Oceano	0.040	0.039	0.038	0.032	0.034	0.026	0.115	0.110	0.105	0.098	0.098	0.084	2.717	2.686	2.634	2.543	2.556	2.408
	Floresta	0.229	0.262	0.288	0.316	0.342	0.377	0.315	0.357	0.383	0.407	0.436	0.466	6.557	7.320	7.758	8.290	8.692	8.933
	Urbana	0.081	0.074	0.063	0.055	0.053	0.034	0.164	0.157	0.143	0.136	0.127	0.099	3.371	3.199	3.010	2.895	2.817	2.435
HV	Oceano	0.012	0.009	0.007	0.008	0.006	0.004	0.049	0.049	0.041	0.039	0.035	0.027	1.971	1.931	1.864	1.837	1.779	1.659
	Floresta	0.009	0.005	0.005	0.004	0.003	0.001	0.043	0.035	0.027	0.024	0.020	0.009	1.833	1.753	1.637	1.589	1.509	1.216
	Urbana	0.083	0.070	0.065	0.059	0.057	0.030	0.170	0.157	0.144	0.134	0.127	0.096	3.419	3.188	3.063	2.939	2.800	2.372
VV	Oceano	0.015	0.015	0.011	0.011	0.011	0.008	0.061	0.059	0.056	0.050	0.048	0.040	2.100	2.086	2.043	1.957	1.931	1.840
	Floresta	0.035	0.032	0.026	0.024	0.018	0.009	0.100	0.097	0.083	0.079	0.070	0.040	2.497	2.480	2.355	2.285	2.157	1.813
	Urbana	0.064	0.053	0.052	0.037	0.033	0.021	0.142	0.127	0.118	0.101	0.089	0.064	3.021	2.821	2.699	2.482	2.364	1.974

Tabela 4.15: Taxas de rejeição para testes de divergência- (h, ϕ) sob $H_0: (\theta_1, L_1, \mu_1) = (\theta_2, L_2, \mu_2)$ e tamanho amostral N para distribuições Γ baseada nos dados

Canal	*Região	nível nominal 1%						nível nominal 5%						\bar{S}_\bullet					
		64	81	100	121	144	225	64	81	100	121	144	225	64	81	100	121	144	225
HH	Oceano	0.049	0.046	0.045	0.037	0.039	0.029	0.124	0.120	0.112	0.104	0.106	0.088	2.823	2.780	2.709	2.607	2.621	2.449
	Floresta	0.324	0.371	0.403	0.431	0.468	0.514	0.398	0.447	0.472	0.501	0.532	0.563	7.151	8.273	9.394	10.616	1.949	14.177
	Urbana	0.255	0.260	0.247	0.249	0.252	0.226	0.367	0.366	0.355	0.360	0.362	0.334	6.823	6.843	6.668	6.582	6.570	5.901
HV	Oceano	0.011	0.010	0.008	0.008	0.006	0.005	0.051	0.048	0.041	0.039	0.037	0.028	1.969	1.933	1.867	1.841	1.784	1.669
	Floresta	0.018	0.013	0.011	0.009	0.007	0.003	0.064	0.054	0.043	0.038	0.034	0.017	2.049	1.960	1.828	1.759	1.679	1.360
	Urbana	0.225	0.222	0.213	0.214	0.205	0.174	0.332	0.321	0.320	0.326	0.312	0.289	6.131	6.040	5.844	5.799	5.572	4.895
VV	Oceano	0.016	0.016	0.013	0.012	0.011	0.009	0.062	0.059	0.058	0.051	0.051	0.043	2.112	2.097	2.057	1.971	1.942	1.853
	Floresta	0.104	0.107	0.090	0.088	0.075	0.044	0.183	0.193	0.184	0.174	0.162	0.121	3.542	3.583	3.394	3.293	3.103	2.590
	Urbana	0.218	0.206	0.190	0.182	0.169	0.135	0.329	0.315	0.294	0.288	0.274	0.231	6.019	5.756	5.530	5.171	4.918	4.242

Considerações finais

Em geral, esta dissertação confirma o conhecimento esperado de que “o modelo $G\Gamma$ supera a lei Γ em termos de medidas de bondade de ajuste.” Conclusão similar também foi obtida por Li *et al.* (2010) e Li *et al.* (2011) sob cenários de intensidade SAR. Aqueles trabalhos adotaram a suposição $G\Gamma$ como uma alternativa empírica para descrever intensidades de pixels de imagens SAR; entretanto, não consideram a formação física destas imagens. Nesta dissertação, consideramos a distribuição $G\Gamma$ como uma suposição estocástica para o ruído *speckle* que estende o modelo clássico Γ com um parâmetro adicional, θ . Este último pode ser interpretado como um fator de correção para o número de looks na presença de diferentes graus de texturas. Esta abordagem é similar àquela em se entender o modelo $G\Gamma$ como um retorno da modelagem multiplicativa tal que a distribuição $G\Gamma$ descreve o ruído, enquanto a distribuição de Dirac [<http://mathworld.wolfram.com/DeltaFunction.html>] é usada para descrever a forma do relevo. Nesta perspectiva, os modelos Γ e $G\Gamma$ se mostram encaixados tanto analítica como fisicamente. Assim, fazia-se necessário um estudo comparativo entre eles. Através de experimentos nos âmbitos de Inferência Estatística, Teoria Assintótica e Teoria da Informação esta dissertação apresentou evidências que a lei $G\Gamma$ pode ser uma alternativa eficiente para a modelagem de dados de intensidade SAR.

A partir da abordagem adotada para descrever intensidades SAR pelo modelo $G\Gamma$, propôs-se um novo método de estimação para os parâmetros $G\Gamma$. Em particular, estimadores em forma fechada tanto para a média de retorno das intensidades como para o número de *looks* foram derivados. Adicionalmente, com base nos EMVs para o número de *looks* e na decomposição espectral da matriz de coerência, elaborou-se um estimador semi-paramétrico para o número de *looks* que pode fornecer melhor desempenho que outros existentes na literatura SAR, tal como o proposto por Anfinsen *et al.* (2009).

Ainda, um método melhorado de estimação por MV foi proposto para a distribuição

$G\Gamma$ por meio da derivação do viés de segunda ordem de acordo com a expressão proposta por Cox e Snell. Neste ponto, foi concluído que estimadores obtidos por meio da correção de viés podem ser indicados como uma alternativa eficiente para se estimar os parâmetros θ e L da distribuição $G\Gamma$. Estimativas menos viesadas foram obtidas, principalmente em pequenos tamanhos de amostra, colocando as expressões de viés como quantidades importantes no nível de pré-processamento.

Por fim, este trabalho propôs seis testes de hipóteses envolvendo distâncias estocásticas: cinco entre os espaços paramétricos Γ e $G\Gamma$ e um entre o modelo Wishart complexo escalonado e o vetor trivariado não-correlacionado em termos das marginais Γ e $G\Gamma$. Nossas propostas demonstraram satisfazer um objetivo duplo: identificar contraste em intensidades SAR e *Full PolSAR*, quantificando a necessidade de adicionar um parâmetro à distribuição Γ (modelo básico para o ruído *speckle*) e a redundância ao se definir um modelo matricial Wishart quando uma distribuição $G\Gamma$ trivariada não-correlacionada poderia ser adotada. Foram derivadas fórmulas simples para as divergências de Kullback-Leibler e Jensen-Shannon entre os modelos Γ e $G\Gamma$ e a divergência de Jensen-Shannon entre $G\Gamma$. Introduzimos também uma quantidade de informação quase mútua entre os modelos Wishart complexo escalonado e $G\Gamma$ trivariado não-correlacionado. Esta última medida foi aplicada para verificar se a magnitude da correlação complexa entre os canais de polarização é estatisticamente nula. Com base nos resultados desta etapa, concluiu-se que as medidas de contraste estocástico entre modelos $G\Gamma$ são melhores do que aquelas entre modelos Γ . Observou-se ainda que as medidas propostas podem ser sensíveis para detectar mudança de contraste, sendo úteis para resolver problemas de detecção de bordas e segmentação. Em termos de comparação entre os modelos Γ e $G\Gamma$ ($D_{KL}(Z||Y)$, $D_{KL}(Y||Z)$), concluímos que a suposição $G\Gamma$ é recomendada para dados com alta variabilidade. Enquanto que o modelo Γ é mais ajustável para dados homogêneos.

Finalmente, o contraste entre o vetor trivariado $G\Gamma(\Gamma)$ não correlacionado e o modelo Wishart complexo escalonado foi empregado buscando a simplificação da suposição da matriz na análise de dados *Full PolSAR*. Como esperado, valores da estatística proposta foram superiores aos limites definidos em termos de percentis qui-quadrado. No entanto, a presença de correlação nula entre os canais de polarização se mostrou mais provável diante de cenários com texturas intensas; por exemplo, em cenas urbanas.

Referências

- ALI, S. M.; SILVEY, S. D. A general class of coefficients of divergence of one distribution from another. *Journal of the Royal Statistical Society. Series B (Methodological)*, Blackwell Publishing for the Royal Statistical Society, v. 28, p. 131–142, 1966.
- ANFENSEN, S. N.; DOULGERIS, A. P.; ELTOFT, T. Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, v. 47, p. 3795–3809, 2009.
- BARTLETT, M. Approximate confidence intervals i. *Biometrika*, v. 40, p. 12–19, 1953.
- BEAULIEU, J.-M.; TOUZI, R. Segmentation of textured polarimetric SAR scenes by likelihood approximation. *IEEE Transactions on Geoscience and Remote Sensing*, v. 42, p. 2063–2072, 2004.
- BLACKNELL, D. Comparison of parameter estimators for k-distribution. *IEE Proceedings-Radar, Sonar and Navigation*, v. 141, p. 45–52, 1994.
- BLATT, D.; III, A. O. H. On tests for global maximum of the log-likelihood function. *IEEE Transactions on Information Theory*, v. 53, p. 2510–2525, 2007.
- CASELLA, G.; BERGER, R. L. *Statistical Inference*. [S.l.]: Duxbury Press, 2002.
- CHAMBERS, J. M. *Software for Data Analysis: Programming with R*. [S.l.]: Springer, 2008.
- CLOUDE, S. Group theory and polarization algebra. *Optik*, v. 75, n. 1, p. 26–36, 1986.
- CLOUDE, S.; POTTIER, E. A. A review of target decomposition theorems in radar polarimetry. *IEEE Transactions on Geoscience and Remote Sensing*, v. 34, n. 2, p. 498–518, 1996.
- CORDEIRO, G. M. Bias correction. In: LOVRIC, M. (Ed.). *International Encyclopedia of Statistical Science*. [S.l.]: Springer Berlin Heidelberg, 2011. p. 148–152.
- COVER, T. M.; THOMAS, J. A. *Elements of Information Theory*. New York: Wiley-Interscience, 1991.

- COX, D.; HINKLEY, D. *Theoretical Statistics*. [S.l.]: Chapman and Hall, 1979. ISBN 9780412161605.
- COX, D. R.; SNELL, E. J. A general definition of residuals. *Journal of the Royal Statistical Society. Series B (Methodological)*, v. 30, p. 248–275, 1968.
- CSISZÁR, I. Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*, v. 2, p. 299–318, 1967.
- DONOHO, D. L.; VETTERLI, M.; DEVORE, R. A.; DAUBECHIES, I. Data compression and harmonic analysis. *IEEE Transactions on Information Theory*, v. 44, p. 2435–2476, 1998.
- DOURADO, G. B. *Correção de Viés do Estimador de Máxima Verossimilhança para a Família Exponencial Biparamétrica*. 2004.
- EGUCHI, S.; COPAS, J. Interpreting Kullback-Leibler divergence with the Neyman-Pearson lemma. *Journal of Multivariate Analysis*, v. 97, p. 2034–2040, 2006.
- FREITAS, C.; SANT’ANNA, S.; RENNÓ, C.; CORREIA, A. Utilização de imagens de radar de abertura sintética na classificação de uso e ocupação do solo. *INPE, São José dos Campos-SP*, 2007.
- FREITAS, C. C.; FRERY, A. C.; CORREIA, A. H. The polarimetric G distribution for SAR data analysis. *Environmetrics*, v. 16, p. 13–31, 2005.
- FRERY, A. C.; ALMEIDA, E. S. de; ROSSO, O. A. The generalized statistical complexity of PolSAR data. In: *Conference on Synthetic Aperture Radar (AP SAR), 2013 Asia-Pacific*. [S.l.: s.n.], 2013. p. 100–103.
- FRERY, A. C.; CINTRA, R. J.; NASCIMENTO, A. D. C. Entropy-based statistical analysis of PolSAR data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 51, p. 3733–3743, 2013.
- FRERY, A. C.; MULLER, H. J.; YANASSE, C. C. F.; SANT’ANNA, S. J. S. A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*, v. 35, p. 648–659, 1997.
- FRERY, A. C.; NASCIMENTO, A. D. C.; CINTRA, R. J. Analytic expressions for stochastic distances between relaxed complex Wishart distribution. *IEEE Transacition on Geoscience and Remote Sensing*, v. 52, p. 1213–1226, 2014.
- GAO, G. Statistical modeling of SAR images: A survey. *Sensors*, v. 10, p. 775–795, 2010.
- GIERULL, C. H.; SIKANETA, I. C. Estimating the effective number of looks in interferometric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 40, p. 1733–1742, 2002.
- GOMES, O.; COMBES, C.; DUSSAUCHOY, A. Parameter estimation of the generalized gamma distribution. *Mathematics and Computers in Simulation*, v. 79, p. 955–963, 2008.

- GOODMAN, N. R. Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction). *The Annals of Mathematical Statistics*, The Institute of Mathematical Statistics, v. 34, p. 152–177, 1963.
- GOUDAIL, F.; RÉFRÉGIER, P.; DELYON, G. Bhattacharyya distance as a contrast parameter for statistical processing of noisy optical images. *Journal of the Optical Society of America A*, v. 21, p. 1231–1240, 2004.
- HAGER, H. W.; BAIN, L. J. Inferential procedures for the generalized gamma distribution. *Journal of the American Statistical Association*, v. 65, p. 1601–1609, 1970.
- HARTLEY, R. V. L. Transmission of information. *Bell System Technical Journal*, v. 7, p. 535–563, 1928.
- JAGER, L.; WELLNER, J. A. Goodness-of-fit tests via phi-divergences. *The Annals of Statistics*, v. 35, p. 2018–2053, 2007.
- JENSEN, J. *Remote Sensing of the Environment: An Earth Resource Perspective*. [S.l.]: Pearson Prentice Hall, 2007.
- KERSTEN, P. R.; LEE, J.-S.; AINSWORTH, T. L. Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering. *IEEE Transactions on Geoscience and Remote Sensing*, v. 43, p. 519–527, 2005.
- KULLBACK, S. *Information Theory and Statistics*. [S.l.]: Dover Publications, 1968.
- KULLBACK, S.; LEIBLER, R. A. On information and sufficiency. *The Annals of Mathematical Statistics*, v. 22, p. 79–86, 1951.
- KUPPERMAN, M. *Further Applications of Information Theory to Multivariate Analysis and Statistical Inference*. Tese (Doutorado), 1957.
- LEE, J. S.; POTTIER, E. *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton: CRC, 2009.
- LI, H.-C.; HONG, W.; WU, Y.-R.; FAN, P.-Z. An efficient and flexible statistical model based on generalized gamma distribution for amplitude SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, v. 48, p. 2711–2722, 2010.
- LI, H.-C.; HONG, W.; WU, Y.-R.; FAN, P.-Z. On the empirical-statistical modeling of SAR images with generalized gamma distribution. *IEEE Journal of Selected Topics in Signal Processing*, v. 5, p. 386–397, 2011.
- LOPES, E. A. C. *Correção de Viés no Modelo de Regressão Normal Assimétrico*. 2007.
- MENÉNDEZ, M. L.; PARDO, J. A.; PARDO, L.; PARDO, M. C. The Jensen-Shannon divergence. *Journal of the Franklin Institute*, v. 334, p. 307–318, 1997.
- MORIO, J.; RÉFRÉGIER, P.; GOUDAIL, F.; DUBOIS-FERNANDEZ, P.; DUPUIS, X. A characterization of Shannon entropy and Bhattacharyya measure of contrast in polarimetric and interferometric SAR image. *Proceedings of the IEEE*, v. 97, p. 1097–1108, 2009.

- NASCIMENTO, A. D. C. *Teoria Estatística da Informação para Dados de Radar de Abertura Sintética Univariados e Polarimétricos*. Tese (Doutorado) — Universidade Federal de Pernambuco, 2012.
- NASCIMENTO, A. D. C.; CINTRA, R. J.; FRERY, A. C. Hypothesis testing in speckled data with stochastic distances. *IEEE Transactions on Geoscience and Remote Sensing*, v. 48, p. 373–385, 2010.
- NASCIMENTO, A. D. C.; HORTA, M. M.; FRERY, A. C.; CINTRA, R. J. Comparing edge detection methods based on stochastic entropies and distances for PolSAR imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 7, p. 648–663, 2014.
- OLIVER, C.; QUEGAN, S. *Understanding Synthetic Aperture Radar Images*. [S.l.]: SciTech Publishing, 1998.
- SALICRÚ, M.; MENÉNDEZ, M. L.; PARDO, L.; MORALES, D. On the applications of divergence type measures in testing statistical hypothesis. *Journal of Multivariate Analysis*, v. 51, p. 372–391, 1994.
- SHANNON, C. E. A mathematical theory of communication. *Bell System Technical Journal*, v. 27, p. 379–423, 1948.
- STACY, E. W. A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 33, p. 1187–1192, 1962.
- TOMA, A.; LEONI-AUBIN, S. Robust tests based on dual divergence estimators and saddlepoint approximations. *Journal of Multivariate Analysis*, v. 101, p. 1143–1155, 2009.
- TORRES, L.; SANT’ANNA, S. J. S.; FREITAS, C. C.; FRERY, A. C. Speckle reduction in polarimetric SAR imagery with stochastic distances and nonlocal means. *Pattern Recognition*, v. 47, p. 141–157, 2014.
- ULABY, F. T.; ELACHI, C. *Radar Polarimetry for Geoscience Applications*. Norwood: Artech House, 1990.
- VASCONCELLOS, K. L. P.; FRERY, A. C.; SILVA, L. B. Improving estimation in speckled imagery. *Computational Statistics*, v. 20, p. 503–519, 2005.