



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

ALISSON DOS SANTOS SILVA

**TEORIA DA INFORMAÇÃO APLICADA A DISTRIBUIÇÕES CONJUNTAS
INDUZIDAS DO ATRIBUTO SPAN (TOTAL SCATTERING POWER IMAGE)
EM IMAGENS POLSAR**

Recife

2023

ALISSON DOS SANTOS SILVA

**TEORIA DA INFORMAÇÃO APLICADA A DISTRIBUIÇÕES CONJUNTAS
INDUZIDAS DO ATRIBUTO SPAN (TOTAL SCATTERING POWER IMAGE)
EM IMAGENS POLSAR**

Tese de Doutorado apresentado ao Programa de Pós-Graduação em Estatística do Departamento de Estatística da Universidade Federal de Pernambuco (UFPE) como requisito para obtenção do título de Doutor em Estatística. Área de concentração: Estatística Aplicada

Orientador: Abraão David Costa do Nascimento

Recife

2023

Catálogo na fonte
Bibliotecária Nataly Soares Leite Moro, CRB4-1722

S586t Silva, Alisson dos Santos
Teoria da informação aplicada a distribuições conjuntas induzidas do atributo SPAN (total scattering power image) em imagens PolSAR / Alisson dos Santos Silva. – 2023.
104 f.: il., fig., tab.

Orientador: Abraão David Costa do Nascimento.
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN, Estatística, Recife, 2023.
Inclui referências e apêndices.

1. Estatística aplicada. 2. Sensoriamento remoto. 3. Divergência. 4. Testes de hipótese. 5. Distribuição bivariada. 6. Modelo multiplicativo PolSAR I. Nascimento, Abraão David Costa do (orientador). II. Título.

310 CDD (23. ed.) UFPE- CCEN 2023 - 78

ALISSON DOS SANTOS SILVA

**TEORIA DA INFORMAÇÃO APLICADA A DISTRIBUIÇÕES CONJUNTAS INDUZIDAS
DO ATRIBUTO SPAN(TOTAL SCATTERING POWER IMAGE) EM IMAGENS POLSAR**

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

Aprovado em: 27 de fevereiro de 2023.

BANCA EXAMINADORA

Prof. Dr. Abraão David Costa do Nascimento
Presidente

Prof. Dr. Alejandro Cesar Frery Orgambide
Examinador Interno

Prof. Dr. Getúlio José Amorim do Amaral
Examinador Interno

Prof. Dr. Anderson A. de Borba
Examinador Externo

Prof. Dr. Jodavid de Araújo Ferreira
Examinador Externo

AGRADECIMENTOS

Agradeço a Deus, pela vida, sabedoria, força, coragem e por guiar todos os meus passos, iluminando os meus caminhos e mostrando sempre a direção certa ao longo da minha vida.

Ao meu orientador, Dr. Prof. Abraão David Costa do Nascimento, pela orientação, sabedoria e paciência durante os quatro anos do doutorado, pela dedicação à elaboração desta tese e dos artigos científicos.

Aos meus pais, Antônio e Maria de Fátima, por todo amor, carinho, apoio e dedicação ao longo da minha vida, incentivando e me apoiando em todos os momentos, ao longo da minha formação acadêmica e profissional.

À minha eterna avó Maria Guedes, para quem destino meu amor e carinho por sua eterna memória.

À minha família, que me apoia em todos os momentos da minha vida.

Aos meus amigos de Recife e aos colegas da pós-graduação, por estarem próximos nos momentos de estudo e de descontração: Adenice, Alexsandra, Anne, Anabeth, Cristina, Codjo, Jodavid, Luan, Lucas, Pedro, além de muitos outros. E aos meus amigos da UFPB: Maizza, Luan Rodrigues e Raul.

Aos professores do PPGE, que foram tão importantes na minha formação acadêmica, por engrandecer meus conhecimentos ao longo desses quatro anos.

À FACEPE, pelo apoio financeiro.

RESUMO

O sistema de radar de abertura sintética polarimétrico (PolSAR) é uma das ferramentas mais utilizadas para resolver problemas de sensoriamento remoto. Tal como em todas as técnicas de imageamento que utilizam iluminação coerente, as imagens PolSAR são contaminadas por um ruído multidimensional granular denominado como 'speckle', que impõe uma natureza não normal e multiplicativa a alguns de seus dados resultantes. Portanto, é necessário um processamento cuidadosamente planejado das imagens PolSAR, como no desenvolvimento de novos testes de hipóteses e detectores de mudança. Nesta tese, usamos a distribuição gama bivariada (MBG) proposta por McKay para descrever distribuições empíricas de dados resultantes de dois componentes de um dos mais importantes atributos no tratamento com dados PolSAR, o total scattered power conhecido como SPAN. Ainda no contexto de uma modelagem especializada, propõe-se, a partir da MBG e da abordagem de modelagem multiplicativa, uma nova distribuição bivariada-chamada $\mathcal{G}_I^0$ McKay bivariada ($\mathcal{G}_I^0\text{MB}$) para descrever um par aleatório induzido do SPAN. Algumas propriedades matemáticas associadas a $\mathcal{G}_I^0\text{MB}$ são derivadas, como vetor de médias, matriz de covariâncias, densidades marginais, função geradora de momentos, esperanças condicionais, variâncias condicionais e momentos condicionais. O objetivo de investigação aplicado nesta tese é construir novos detectores de mudança com base em novos testes de hipóteses formulados das distribuições induzidas pelo SPAN, a saber MBG e $\mathcal{G}_I^0\text{MB}$. Para esse fim, derivam-se expressões em forma fechada para as divergências de Kullback-Leibler e Rényi a partir das distribuições MBG e $\mathcal{G}_I^0\text{MB}$. Como consequência, novos testes de hipótese para duas amostras são introduzidos e seus desempenhos são analisados via experimentos de Monte Carlo. Os resultados de simulação indicam que testes baseados em divergência podem ter um desempenho superior ao teste da razão entre verossimilhança, tomando a estimativa da probabilidade do erro tipo I como critério de avaliação. Finalmente, os novos testes foram aplicados às imagens reais PolSAR para avaliar as mudanças causadas por processos de urbanização em regiões de Los Angeles-Califórnia. Os resultados indicam que as propostas podem detectar mudanças entre as imagens PolSAR com desempenho de critérios de avaliação melhor do que a detecção feita com base na razão entre verossimilhanças.

Palavras-chave: sensoriamento remoto; divergência; testes de hipótese; distribuição bivariada; modelo multiplicativo PolSAR.

ABSTRACT

Polarimetric synthetic aperture radar (PolSAR) is one of the most widely used tools for solving remote sensing problems. As with all imaging techniques that use coherent illumination, PolSAR images are contaminated by a multidimensional granular noise called speckle, which imposes a non-normal, multiplicative nature on some of their resulting data. Therefore, carefully planned processing of PolSAR images is necessary, for example in the development of new hypothesis tests and change detectors. In this thesis, we use the bivariate gamma distribution (MBG) proposed by McKay to describe empirical data distributions resulting from two components of one of the most important attributes in dealing with PolSAR data, the total scattered power known as SPAN. Still in the context of expert modeling, a new bivariate distribution called $\mathcal{G}_I^0$ McKay bivariate ($\mathcal{G}_I^0$ MB) is proposed (from MBG and the multiplicative modeling approach) to describe an induced SPAN random pair. Some mathematical properties associated with $\mathcal{G}_I^0$ MB are derived: such as mean vector, covariance matrix, marginal densities, moment generating function, conditional expectancies, conditional variances, and conditional moments. The research objective applied in this thesis is to construct new change detectors based on new hypothesis tests formulated from the SPAN-induced distributions, namely MBG and $\mathcal{G}_I^0$ MB. For this purpose, closed-form expressions are derived for Kullback-Leibler and Rényi divergences from MBG and $\mathcal{G}_I^0$ MB distributions. As a consequence, new hypothesis tests for two samples are introduced and their performances are analyzed via Monte Carlo experiments. The simulation results indicate that divergence-based tests can outperform the likelihood ratio test by taking the type I error probability estimate as the evaluation criterion. Finally, the new tests were applied to real PolSAR images to evaluate changes caused by urbanization processes in Los Angeles-California regions. The results indicate that the proposals can detect changes among PolSAR images with better evaluation criteria performance than the detection done based on likelihood ratio.

Keyword: remote sensing; divergence; hypothesis testing; bivariate distribution; multiplicative model PolSAR.

LISTA DE FIGURAS

Figura 1 – Descrição do Sistema SAR.	15
Figura 2 – Diagrama de blocos para testes de hipóteses e detectores de mudança baseados nas divergências de Kullback-Leibler e Rényi. $\chi_{m,\eta}^2$, em que η representa quantil.	29
Figura 3 – Ilustração de alvos associados com valores diferentes do parâmetro de rugosidade α de uma distribuição $\mathcal{G}_m^0$	31
Figura 4 – Medidas de distância entre duas variáveis aleatórias MBG em função de α_1	38
Figura 5 – Cenários de simulação obtida do campo de ensaio agrícola de Foulum em Jutland (Dinamarca), gráfico de dispersão de valores de pares (HH, HH+VV), e curvas de nível associadas.	41
Figura 6 – Mapas de correlação MGB para as imagens de Los Angeles (USA), capturadas em 2015.	42
Figura 7 – Mapas de detecção de mudanças do SPAN HH-HV para região de Los Angeles (CA, USA).	48
Figura 8 – Mapas de detecção de mudanças do SPAN HH-VV e HH-VV para região de Los Angeles (CA, USA).	49
Figura 9 – Mapas de detecção de mudanças do SPAN HH-HV para região de Los Angeles (CA, USA).	50
Figura 10 – Mapas de detecção de mudanças do SPAN HH-VV e HH-VV para região de Los Angeles (CA, USA).	51
Figura 11 – Cenários de simulação obtida da região de São Francisco (USA) e gráfico de dispersão de valores de pares (HH, HH+VV).	59
Figura 12 – Curvas de níveis teóricas e empíricas ajustadas a fdp $\mathcal{G}_I^0$ MB.	60
Figura 13 – Medidas de distância entre duas variáveis aleatórias $\mathcal{G}_I^0$ MB em função de α_1	66
Figura 14 – Mapas de detecção de mudanças do SPAN HH-HV para região de Los Angeles (CA, USA).	75
Figura 15 – Mapas de detecção de mudanças do SPAN HH-VV e HV-VV para região de Los Angeles (CA, USA).	76
Figura 16 – Mapas de detecção de mudanças do SPAN HH-HV para região de Los Angeles (CA, USA).	77

Figura 17 – Mapas de detecção de mudanças do SPAN HH-VV e HV-VV para região de Los Angeles (CA, USA).	78
---	----

LISTA DE TABELAS

Tabela 1 – Valor dos parâmetros para os cenários simulados.	40
Tabela 2 – Taxas de Rejeição de Testes de Divergência sob $H_0 : (\alpha_{11}, \alpha_{12}, \gamma_1) =$ $(\alpha_{21}, \alpha_{22}, \gamma_2)$	43
Tabela 3 – Desempenho dos detectores - Cena 1 - Los Angeles (CA, USA).	46
Tabela 4 – Desempenho dos detectores - Cena 2 - Los Angeles (CA, USA).	47
Tabela 5 – Taxas de Rejeição de Testes de Divergência sob $H_0 : (\alpha_{11}, \alpha_{12}, \gamma_1) =$ $(\alpha_{21}, \alpha_{22}, \gamma_2)$ when $\beta = 1.5$	69
Tabela 6 – Taxas de Rejeição de Testes de Divergência sob $H_0 : (\alpha_{11}, \alpha_{12}, \gamma_1) =$ $(\alpha_{21}, \alpha_{22}, \gamma_2)$ when $\beta = 2$	70
Tabela 7 – Taxas de Rejeição de Testes de Divergência sob $H_0 : (\alpha_{11}, \alpha_{12}, \gamma_1) =$ $(\alpha_{21}, \alpha_{22}, \gamma_2)$ when $\beta = 2.5$	71
Tabela 8 – Desempenho dos detectores - Cena 1 - Los Angeles (CA, USA).	73
Tabela 9 – Desempenho dos detectores - Cena 2 - Los Angeles (CA, USA).	74

LISTA DE SÍMBOLOS

$\Gamma(\cdot)$	Função Gama
$\beta(\cdot)$	Função Beta
$\mathbb{E}(\cdot)$	Esperança
$\mathbb{V}\text{ar}(\cdot)$	Variância
$\text{Cor}(\cdot)$	Correlação
$\psi(\cdot)$	Função Psi
$\top$	Transposto
$\mathcal{D}$	Convergência em distribuição.
χ_p^2	Distribuição Qui-Quadrado
${}_pF_q(\cdot)$	Função Hipergeométrica
HH	Canal de intensidade HH de uma imagem SAR/PoLSAR
HV	Canal de intensidade HV de uma imagem SAR/PoLSAR
VH	Canal de intensidade VH de uma imagem SAR/PoLSAR
VV	Canal de intensidade VV de uma imagem SAR/PoLSAR
$\mathcal{G}_I^0\text{MB}$	$\mathcal{G}_I^0$ McKay bivariada
$D_{\text{KL}}(\cdot \parallel \cdot)$	Divergência de Kullback-Leibler
$D_{\text{R:r}}(\cdot \parallel \cdot)$	Divergência de Rényi
$d_{\text{KL}}(\cdot \parallel \cdot)$	Distância de Kullback-Leibler
$d_{\text{R:r}}(\cdot \parallel \cdot)$	Distância de Rényi
$S_{KL}(\cdot, \cdot)$	Estatística baseada na Divergência de Kullback-Leibler
$S_R^r(\cdot, \cdot)$	Estatística baseada na Divergência de Rényi
$S_{RV}(\cdot, \cdot)$	Estatística baseada no teste da razão entre verossimilhanças

CD_{KL}	Detector de mudança baseado na distância Kullback-Leibler
CD_R^α	Detector de mudança baseado na distância Rényi de ordem α
CD_{RV}	Detector de mudança baseado na estatística da razão entre verossimilhanças

SUMÁRIO

1	INTRODUÇÃO	14
1.1	OBJETIVOS GERAIS	18
1.1.1	Objetivos Específicos	18
1.2	CONTRIBUIÇÕES	19
1.3	PESQUISAS REALIZADAS	20
1.4	ORGANIZAÇÃO DO TEXTO	20
2	MÉTODOS APLICADOS PARA DADOS POLSAR	22
2.1	FUNDAMENTOS DA TEORIA DA INFORMAÇÃO ESTATÍSTICA	22
2.2	ATRIBUTO POLSAR: SPAN	26
2.3	DETECTORES DE MUDANÇAS POLSAR	27
3	DISTRIBUIÇÃO BIVARIADA MCKAY APLICADA ÀS CARACTERÍSTICAS SAR BASEADAS EM SPAN E MODELAGEM MULTIPLICATIVA	30
3.1	TESTE DA RAZÃO DE VEROSSIMILHANÇA PARA A DISTRIBUIÇÃO GAMA BIVARIADA	34
3.2	NOVOS TESTES DE HIPÓTESE BASEADOS EM DIVERGÊNCIAS PARA A DISTRIBUIÇÃO GAMA BIVARIADA	36
3.3	ESTUDO DE SIMULAÇÃO	39
3.4	APLICAÇÃO EM DADOS REAIS	44
4	DISTRIBUIÇÃO BIVARIADA $\mathcal{G}_I^0\text{MB}$ APLICADA ÀS CARACTERÍSTICAS SAR COM BASE EM SPAN	52
4.1	ESTIMATIVA DE MÁXIMA VEROSSIMILHANÇA PARA OS PARÂMETROS $\mathcal{G}_I^0\text{MB}$	55
4.2	ESPERANÇA E MOMENTOS CONDICIONAIS	56
4.3	APLICAÇÃO A DADOS POLSAR COM BASE NO SPAN PARA $\mathcal{G}_I^0\text{MB}$	58
4.4	TESTE DE HIPÓTESE BASEADOS NA ESTATÍSTICA DA RAZÃO ENTRE VEROSSIMILHANÇAS PARA A $\mathcal{G}_I^0\text{MB}$	61
4.5	NOVOS TESTES DE HIPÓTESES BASEADOS EM DIVERGÊNCIAS PARA A DISTRIBUIÇÃO BIVARIADA $\mathcal{G}_I^0\text{MB}$	63
4.6	ESTUDO DE SIMULAÇÃO	67

4.7	APLICAÇÃO EM DADOS REAIS	72
5	CONCLUSÃO	79
5.1	FUTUROS TRABALHOS	81
	REFERÊNCIAS	82
	APÊNDICE A – DIVERGÊNCIA DE KULLBACK-LEIBLER PARA DISTRIBUIÇÃO MCKAY	87
	APÊNDICE B – DIVERGÊNCIA DE RÉNYI DE ORDEM r PARA DISTRIBUIÇÃO MCKAY	90
	APÊNDICE C – NOVA DISTRIBUIÇÃO BIVARIADA PARA DA- DOS SPECKLED	91
	APÊNDICE D – FUNÇÕES MARGINAIS E FUNÇÃO DE COR- RELAÇÃO PARA $\mathcal{G}_I^0\text{MB}$	92
	APÊNDICE E – FUNÇÃO GERATRIZ DE MOMENTOS PARA $\mathcal{G}_I^0\text{MB}$	94
	APÊNDICE F – ESPERANÇA, VARIÂNCIA E MOMENTOS $\mathcal{G}_I^0\text{MB}$ DO TIPO $Z_1 Z_2$	96
	APÊNDICE G – ESPERANÇA, VARIÂNCIA E MOMENTOS $\mathcal{G}_I^0\text{MB}$ DO TIPO $Z_2 Z_1$	98
	APÊNDICE H – DIVERGÊNCIA DE KULLBACK-LEIBLER PARA DISTRIBUIÇÃO BIVARIADA $\mathcal{G}_I^0\text{MB}$	100
	APÊNDICE I – DIVERGÊNCIA DE RÉNYI DE ORDEM r PARA DISTRIBUIÇÃO BIVARIADA $\mathcal{G}_I^0\text{MB}$	104

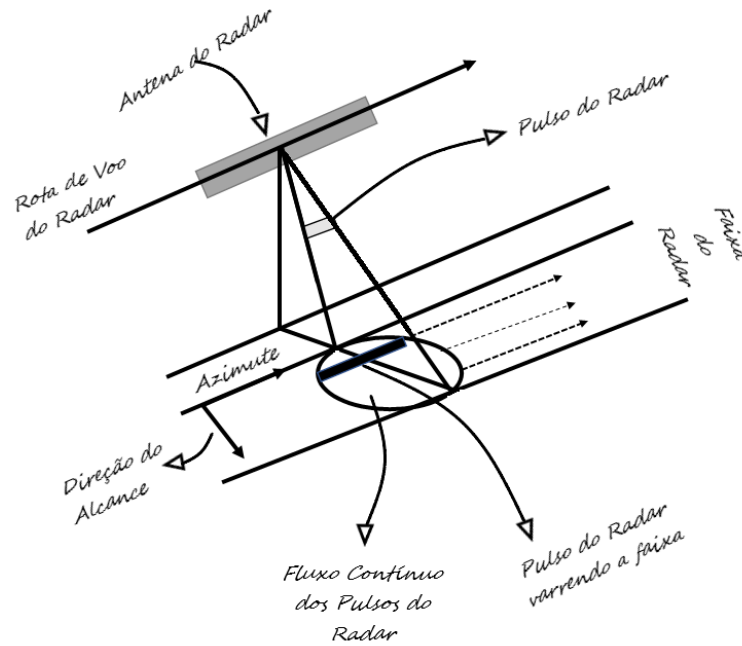
1 INTRODUÇÃO

O radar de abertura sintética (Synthetic Aperture Radar - SAR) usa um mecanismo de iluminação coerente do solo, transmitindo sinais eletromagnéticos de micro-ondas de antenas ligadas aos sensores espaciais. Esses sensores são operados a partir de plataformas elevadas, aeronaves tripuladas e não tripuladas de todos os tipos e satélites que orbitam sobre a superfície terrestre. Alguns sistemas SAR são descritos a seguir:

- O laboratório de propulsão a jato da NASA (JPL) desenvolveu um sistema SAR polarimétrico (PolSAR) aerotransportado que funciona com comprimento de onda de 25 cm e com ângulos de incidência de 0° a 60° , operando na banda L para aplicações de detecção remota. Seus impulsos SAR são transmitidos por uma antena de feixe, paralela ao vetor de velocidade do sistema de aeronaves Convair-990 (THOMPSON, 1986).
- O Centro Canadense para Sensoriamento Remoto (CCRS) planejou o sistema SAR aerotransportado chamado de CCRS C/X- SAR. Esse sistema é formado por dois radares que operam nas bandas C e X de multi-frequência e multi-polarização horizontal e vertical com ângulos de incidência de 0° a 80° . Esse sistema é utilizado nos estudos dos oceanos, terra e derretimento de geleiras (LIVINGSTONE et al., 1988).
- O Laboratório Nacional da Sandia (Califórnia) projetou o sistema SAR aerotransportado chamado Twin-Otter SAR, que opera em quatro bandas: Ka, Ku, X e VHF/UHF, com ângulos de incidência de 2° a 90° . Esse sistema foi aplicado para mapear tanto a zona rural como regiões urbanas (WALKER et al., 1996; CHAN; KOO, 2008).
- A Agência espacial argentina desenvolveu o sistema SAOCOM-1 SAR, composto por dois satélites operados sobre a banda L, com ângulos de incidência de 18° a 50° . Ele fornece informações sobre agricultura, hidrologia, incluindo inundações e emergências, gera mapas operacionais de umidade do solo e explora a capacidade interferométrica SAR (GIRALDEZ, 2004).

O sistema SAR detecta elementos de um determinado alvo, tais como, velocidade, amplitude, localização e direção do movimento (HAJNSEK; DESNOS, 2020; ZYL, 2011; YANG et al., 2021). A geometria do SAR é definida em duas etapas: a primeira é responsável por reduzir os efeitos da direção azimute paralela à trajetória de voo do instrumento do radar, e a segunda

Figura 1 – Descrição do Sistema SAR.



Fonte: Elaborada pelo autor (2023)

aborda o impacto da faixa, tendo localização perpendicular (HAJNSEK; DESNOS, 2020). Essas etapas podem ser vistas na figura 1. Portanto, os modos de operação do sistema SAR determinam a geometria da imagem, a resolução do radar, a largura da faixa e o intervalo de alcance do receptor (HAJNSEK; DESNOS, 2020; ZYL, 2011).

Imagens SAR são formadas combinando os sinais emitidos por ambas as dimensões: a direção do alcance (range) e do azimute. As intensidades dos retornos que compõem uma imagem SAR dependem da energia recebida pelo sensor, como a natureza física do objeto alcançado e o ângulo de incidência dos raios no alvo. Essa onda eletromagnética transmitida é polarizada na direção horizontal H e vertical V (HAJNSEK; DESNOS, 2020).

Os dados SAR polarimétricos (Polarimetric SAR -POLSAR) carregam múltiplas informações sobre uma mesma cena representados em termos de uma matriz de espalhamento de polarização. Cada elemento dessa matriz é resultante da interseção de direções dos sinais transmitidos pelo alvo e recebidos pelo sensor, HH , VV , HV e VH (ZYL, 2011):

$$\mathbf{S} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \quad (1.1)$$

Entre os recursos de sensoriamento remoto, os radares de abertura sintética polarimétricos têm recebido destaque por, dentre outras coisas, operarem em várias condições atmosféricas, produzirem imagens em alta resolução espacial e fornecerem múltiplas informações sobre uma mesma cena. As informações adicionais latentes aos retornos (nas formas das matrizes de *coerência* e de *espalhamento*) dessas imagens de radar podem ser exploradas pelos chamados "teoremas de decomposição polarimétrica" para extrair mais atributos dos conjuntos de dados polarimétricos (Ver Capítulo 6 de Lee e Pottier (2009)). Teoremas de decomposição de alvos foram formalizados pela primeira vez por Huynen (1970) e Pottier (1992). Desde então, houve outras decomposições propostas que podem ser categorizadas em quatro tipos principais:

- *Baseada na dicotomia da matriz de Kennaugh*: Decomposições de Huynen (1990), Barnes (1988) e Yang et al. (1998), Yang et al. (2006);
- *Baseada em modelos matriciais a partir da matriz de coerência*: Decomposições de Freeman e Durden (1998) e Yamaguchi et al. (2005), Yamaguchi, Yajima e Yamada (2006);
- *Baseada em autovetores a partir da matriz de coerência*: Decomposições de Cloude e Pottier (1996), Barnes (1988) e Van (1992);
- *Baseada na decomposição da matriz de espalhamento*: Decomposições de Pauli Lee e Pottier (2009), Krogager (1990) e Cameron e Rais (2006).

Uma técnica eficaz para a decomposição das imagens do PolSAR é a decomposição básica de Pauli, que modela a matriz de espalhamento em três estruturas: a superfície áspera (rough surface), o volume (bulk) e o salto duplo (double jump). Essa decomposição fornece uma imagem com informações sobre o poder de dispersão total da imagem (backscattered), chamado SPAN (YANG et al., 2021). A primeira estrutura é gerada pelo recebimento dos sinais eletromagnéticos no solo ou em água translúcida, que é sensível à dispersão do canal VV. Já a segunda estrutura se origina por meio da captação dos sinais de folhas e galhos em um dossel da floresta e é sensível aos canais VH e HV. A última estrutura é causada por edifícios, troncos de árvores ou vegetação inundada e é mais sensível a um sinal HH polarizado (HAJNSEK; DESNOS, 2020).

O SPAN é uma soma das intensidades de todos os canais polarimétricos da matriz (1.1) (GAO, 2019). O SPAN tem sido aplicado em diversas áreas do pós-processamento de imagens PolSAR, como na segmentação e classificação em áreas urbanas heterogêneas (WANG et al.,

2018; WANG et al., 2017) e na detecção de mudanças em áreas urbanas densas e de grande área florestal (ANSARI; BUDDHIRAJU; MALHOTRA, 2020). O pós-processamento é a etapa que ocorre depois da análise da imagem e envolve técnicas para aprimorar os resultados obtidos na análise, como, por exemplo: filtragem, ajuste de contraste e brilho, e realce de bordas. Essas técnicas ajudam a melhorar a qualidade dos resultados obtidos na análise de imagem. (LEE; POTTIER, 2009).

A criação de mapa de informações a partir do SPAN foi utilizado por alguns autores para melhorar o desempenho de classificação, detecção de mudanças e métodos de segmentação em imagens PolSAR. Ansari, Buddhiraju e Malhotra (2020) e Canisius et al. (2018) utilizaram o SPAN no sensoriamento remoto de variações topográficas entre montanhas e áreas urbanas para prevenção contra terremotos. West e Riley (2019) empregaram essa importante quantidade em diferentes regiões com baixo espalhamento devido a sombras ou vegetação do radar. Wang et al. (2018) combinaram o SPAN com a distância do vetor aleatório invariante esférico para áreas urbanas heterogêneas e Xu et al. (2014) usaram o SPAN para derivar o volume espelhado de imagens de áreas urbanas e rurais.

A detecção de mudanças é uma área bastante importante para o pós-processamento de imagens PolSAR, sendo utilizada para monitorar desastres naturais (WASHAYA; BALZ; MOHAMADI, 2018) bem como identificar mudanças sociodemográficas em posições geográficas (NASCIMENTO; FRERY; CINTRA, 2018).

No processamento da imagem, o uso de modelos probabilísticos (principalmente fundamentados em abordagens relacionadas à formação física de imagens SAR) e de áreas da Estatística clássica ou combinadas à Teoria da informação (como inferência, modelos espaço temporais e aprendizagem de máquina) tem um papel fundamental na interpretação de dados extraídos de imagens PolSAR. Os dados de intensidade SAR ou PolSAR multilook costumam ser modelados pelas seguintes distribuições: $\mathcal{G}$ para intensidade do SAR (FRERY et al., 1997), $\mathcal{G}$ e Wishart complexa e escalonada para dados PolSAR (FREITAS; FRERY; CORREIA, 2005) e a gama multivariada para intensidades multivariadas (HAGEDORN et al., 2006).

A utilização de modelos estatísticos no processamento de imagens SAR vem revelando ótimos resultados como: (i) na detecção de estruturas complexas em áreas urbanas extremamente heterogêneas (INCE; KIRANYAZ; GABBOUJ, 2012), (ii) na identificação de alterações em imagens SAR multilook, usando a teoria de Dezert-Smarandache (GAO; QIN; ZHOU, 2013), (iii) para detectar grandes navios em áreas marítimas densas (HACHICHA; CHAABANE, 2009; CORDEIRO et al., 2019), (iv) para modelar a amplitude e a intensidade do SPAN para diferentes

tipologias de cobertura do solo (BORBA; MARENGONI; FRERY, 2019; LI et al., 2010), (v) no desenvolvimento de novas técnicas para segmentação, classificação e detecção de mudanças em dados simulados e reais (LI et al., 2011), dentre outros.

1.1 OBJETIVOS GERAIS

Neste trabalho de Tese de Doutorado novos testes de hipóteses são propostos para atributos bivariados induzidos a partir da matriz de coerência PolSAR, sendo o principal deles o SPAN. Embora o sistema PolSAR seja uma das técnicas de sensoriamento remoto mais utilizadas, as imagens PolSAR são afetadas por um padrão de interferência granular chamado ruído speckle. Então, para explicar o comportamento do ruído e do relevo nos pares de intensidade acumulada baseado em SPAN, $(||S_{HH}||^2, ||S_{HH}||^2 + ||S_{VV}||^2)$, $(||S_{VV}||^2, ||S_{VV}||^2 + ||S_{HH}||^2)$, $(||S_{HV}||^2, ||S_{HV}||^2 + ||S_{VV}||^2)$ e $(||S_{HV}||^2, ||S_{HV}||^2 + ||S_{HH}||^2)$, requer-se uma modelagem especializada.

A suposição distribucional para retornos hermitianos em áreas homogêneas é a Wishart complexa, e portanto a distribuição conjunta de sua diagonal principal é uma gama multivariada. Do melhor que conhecemos, a partir da literatura PolSAR, a distribuição conjunta do SPAN e seus componentes em cenários homogêneos (como também em outros mais texturados) carece de investigação científica.

Neste manuscrito, endereçamos essa lacuna tanto pela Estatística Clássica como pelo uso de Teoria da Informação. Inicialmente, utilizamos a distribuição gama bivariada de McKay para cenários homogêneos e, subsequentemente, propomos uma nova distribuição de probabilidade para outros tipos de terrenos (deduzida a partir da modelagem multiplicativa), denominada $\mathcal{G}_I^0\text{MB}$. Propriedades matemáticas iniciais relativas a esse novo modelo são propostas, como matriz de covariância, esperanças e variâncias condicionais e a função geradora de momentos.

1.1.1 Objetivos Específicos

- Derivação das divergências e distâncias de Kullback-Leibler (KL) e Rényi para pares aleatórios induzidos a partir do SPAN utilizando as distribuições bivariadas Gama MCKay e $\mathcal{G}_I^0\text{MB}$;
- Formulação de novos testes de hipóteses para duas amostras para as distribuições bivari-

adas Gama MCKay e $\mathcal{G}_I^0\text{MB}$. Desenvolvimento de detectores de mudança entre imagens PolSAR;

- Quantificação do desempenho dos testes propostos por experimentos de Monte Carlo, comparativamente ao do teste da Razão entre Verossimilhanças (RV) (CASELLA; BERGER, 2021), comumente adotado no processamento de imagens PolSAR.

1.2 CONTRIBUIÇÕES

No que se segue, resumos das principais contribuições são fornecidos.

✓ PRIMEIRA CONTRIBUIÇÃO - Capítulo 3:

Neste capítulo, utilizou-se a distribuição gama bivariada de McKay (MBG) para descrever distribuições empíricas resultantes de dois componentes a partir do SPAN. Verificou-se a existência de uma relação entre a medida de correlação MBG e o tipo de textura em uma imagem SAR. Derivaram-se expressões em forma fechada para as divergências de Kullback-Leibler e Rényi da lei MBG. Novos testes de hipótese baseados em divergência e na razão de verossimilhanças para duas amostras são propostos e avaliados usando experimentos de Monte Carlo. Finalmente, os novos testes são aplicados a imagens PolSAR reais para avaliar as mudanças causadas por processos de urbanização nas regiões de Los Angeles, Califórnia-EUA. Os resultados apontaram que o teste baseado na divergência de Rényi de alta ordem capturou melhores mudanças próximas às regiões centrais da imagem. O teste baseado na divergência KL capturou melhores mudanças próximas às regiões da borda da imagem.

✓ SEGUNDA CONTRIBUIÇÃO - Capítulo 4:

Neste capítulo, uma nova distribuição bivariada - chamada $\mathcal{G}_I^0\text{MB}$ - é proposta para descrever distribuições empíricas resultantes do SPAN, utilizando a abordagem de modelagem multiplicativa. Algumas propriedades matemáticas da distribuição $\mathcal{G}_I^0\text{MB}$ são deduzidas: vetor de médias, matriz de covariâncias, densidades marginais, função geradora de momentos, esperanças condicionais, variâncias condicionais e momentos condicionais. Como ilustração do uso da lei $\mathcal{G}_I^0\text{MB}$ em imagens PolSAR, trechos de dados reais da imagem PolSAR de São Francisco são modelados usando curvas de nível a partir da estimação por máxima verossimilhança.

✓ TERCEIRA CONTRIBUIÇÃO - Capítulo 4:

Neste capítulo, derivam-se expressões em forma fechada para as divergências Kullback-Leibler e Rényi para a lei $\mathcal{G}_I^0\text{MB}$. Dessas quantidades, novos testes de hipótese baseados em divergência e razão de verossimilhanças para duas amostras são introduzidos. Um estudo sintético é também feito a fim de quantificar o desempenho deles utilizando experimentos de Monte Carlo. Finalmente, aplicamos os novos testes a imagens reais PolSAR para avaliar as mudanças causadas por processos de urbanização, nas regiões de Los Angeles, Califórnia-EUA.

1.3 PESQUISAS REALIZADAS

Alguns resultados desta tese foram publicados em periódicos clássicos nas áreas de Estatística e Sensoriamento Remoto:

- ✓ Divergence-based tests for the bivariate gamma distribution applied to Polarimetric Synthetic Aperture Radar - *Publicado no Statistical Papers*, em 30 agosto 2022. - Acesso em: <https://doi.org/10.1007/s00362-022-01354-4>
- ✓ A novel bivariate distribution and its divergence-based hypothesis inference. An application to the deforestation focus detection - *Publicado no International Journal of Remote Sensing* em 3 Janeiro 2023 - Acesso em: <https://doi.org/10.1080/01431161.2022.2155087>

1.4 ORGANIZAÇÃO DO TEXTO

Em geral, esta Tese está estruturada nos seguintes capítulos:

- O Capítulo 2, apresenta-se uma revisão PolSAR sobre Teoria de Informação conectada à Estatística e ao atributo SPAN.
- O Capítulo 3, se inicia com uma revisão metodológica sobre a Modelagem multiplicativa e a distribuição McKay bivariada (MBG). Subsequentemente, introduz-se um ensaio teórico sobre dedução das divergências e distâncias de KL e Rényi, para a distribuição MBG. Novos testes de hipóteses para duas amostras baseados nas divergências KL, Rényi e na estatística da razão de verossimilhanças são apresentados para a distribuição MBG. Adicionalmente, novos detectores de mudança em imagens PolSAR são desenvolvidos.

- No Capítulo 4, uma nova distribuição bivariada chamada ($\mathcal{G}_I^0\text{MB}$) é proposta. Além disso, algumas propriedades matemáticas são deduzidas: vetor de médias, matriz de covariâncias, densidades marginais, função geradora de momentos, esperanças condicionais, variâncias condicionais e momentos condicionais. Expressões para divergências e distâncias KL e Rényi na distribuição $\mathcal{G}_I^0\text{MB}$ são propostas. Propomos novos testes de hipóteses para duas amostras, baseados nas divergências KL, Rényi e na estatística da razão entre verossimilhanças $\mathcal{G}_I^0\text{MB}$. Novos detectores de mudança para imagens PolSAR são construídos.
- No Capítulo 5, apresenta-se a conclusão deste trabalho e alguns indicadores de pesquisas futuras.

Este trabalho de Tese de doutorado foi escrito usando $\text{\LaTeX}$, um sistema de composição tipográfica que inclui recursos destinados à produção de documentação técnica e científica (LAMPORT, 1986). As avaliações numéricas e aplicações dos Capítulos 3 e 4 foram realizadas utilizando a linguagem de programação R versão 4.1.0, que se encontra disponível gratuitamente em: <http://www.r-project.org>. O sistema operacional utilizado foi Windows 11, processador i5- 8ºger — 64 bits e 8 GB de memória RAM.

2 MÉTODOS APLICADOS PARA DADOS POLSAR

2.1 FUNDAMENTOS DA TEORIA DA INFORMAÇÃO ESTATÍSTICA

A teoria da informação é uma abordagem clássica para medir a quantidade de informação em um sinal ou mensagem. Foi proposta pelo matemático americano Hartley (1928), sendo baseada em uma medida simples de informação, que é proporcional ao logaritmo do número de símbolos possíveis em um sinal, no contexto da comunicação. Porém, ela apresenta algumas limitações, como não considerar a correlação entre os símbolos em um sinal, e não é adequada para medir a complexidade de sistemas de comunicação mais sofisticados.

Portanto, existe uma complementariedade, na prática, entre a teoria da informação e a teoria estatística de Fisher (FISHER, 1922), e, embora as duas teorias difiram em seus objetivos, elas se relacionam para extrair informações úteis de um determinado conjunto de dados. A teoria estatística de Fisher ajusta modelos estatísticos aos dados e faz inferências estatísticas sobre os parâmetros do modelo.

Posteriormente, Shannon (1948) formalizou essa abordagem amplamente utilizada para medir a quantidade de informação contida em um conjunto de dados, através da análise da distribuição de probabilidade e definiu a entropia que representa a incerteza associada à distribuição de probabilidade.

Finalmente, a noção de entropia relativa (que mais tarde foi chamada de divergência) foi introduzida por Kullback e Leibler (1951). A divergência de Kullback-Leibler pode ser entendida como um critério de comparação entre duas distribuições. Em Estatística, ela é definida como o Valor Esperado do logaritmo da razão entre as densidades de probabilidade, em que uma distribuição é a distribuição verdadeira dos dados e a outra é a distribuição estimada por um modelo:

Sejam $\mathbf{x}$ e $\mathbf{y}$ dois vetores aleatórios com funções densidades de probabilidade $f_{\mathbf{x}}(\mathbf{z}; \boldsymbol{\theta}_{\mathbf{x}})$ e $f_{\mathbf{y}}(\mathbf{z}; \boldsymbol{\theta}_{\mathbf{y}})$ e com suporte $\mathcal{Z} \subset \mathbb{R}^p$. A divergência de Kullback-Leibler $\mathbf{x}$ em relação a $\mathbf{y}$ é definida por:

$$D_{\text{KL}}(\mathbf{x} \parallel \mathbf{y}) = \mathbb{E}_{\mathbf{x}} \left[\log \frac{f_{\mathbf{x}}(\mathbf{x}; \boldsymbol{\theta}_{\mathbf{x}})}{f_{\mathbf{y}}(\mathbf{x}; \boldsymbol{\theta}_{\mathbf{y}})} \right] = \int_{\mathcal{Z}} f_{\mathbf{x}}(\mathbf{z}; \boldsymbol{\theta}_{\mathbf{x}}) \log \frac{f_{\mathbf{x}}(\mathbf{z}; \boldsymbol{\theta}_{\mathbf{x}})}{f_{\mathbf{y}}(\mathbf{z}; \boldsymbol{\theta}_{\mathbf{y}})} d\mathbf{z}, \quad (2.1)$$

em que $d\mathbf{z} = \prod_{i=1}^p dz_i$, $\mathbb{E}_{\mathbf{x}}[\mathcal{V}(\mathbf{x})] = \int_{\mathcal{Z}} \mathcal{V}(\mathbf{z}) f_{\mathbf{x}}(\mathbf{z}; \boldsymbol{\theta}_{\mathbf{x}}) d\mathbf{z}$ e $\mathcal{V}(\cdot)$ é uma função mensurável cuja esperança existe. A partir de agora, assumirá $p = 2$, pois o foco desta tese será com pares

aleatórios.

É perceptível que a Equação (2.1) é sempre não-negativa e é zero somente se $f_x(\mathbf{z}; \boldsymbol{\theta}_x) = f_y(\mathbf{z}; \boldsymbol{\theta}_y)$ por todos os $\mathbf{z} \in \mathcal{Z}$. A divergência de Kullback-Leibler (KL), $D_{KL}(\mathbf{x} \parallel \mathbf{y})$ é uma quantidade que expressa a ineficiência de supor a distribuição de $\mathbf{y}$ quando o modelo $\mathbf{x}$ é verdadeiro (COVER; THOMAS, 2007). De fato, tal medida é conhecida sob vários termos, tais como entropia cruzada, divergência de informações e informações para discriminação.

Kullback (1968), Csiszár (1967) e Amari (1985) apresentaram um trabalho sistemático e abrangente sobre a medida de divergência KL. Kullback (1968) mostrou que a KL é uma medida assimétrica de dissimilaridade entre distribuições de probabilidade. Csiszár (1967) mostrou que a KL pode ser vista como um caso especial de uma família mais ampla de medidas de divergência, conhecidas como medidas de informação tipo. Amari (1985) mostrou que a KL pode ser vista como uma medida da distância entre duas distribuições de probabilidade em um espaço de parâmetros, e desenvolveu ferramentas matemáticas para calcular e interpretar a KL nesse contexto.

Embora $D_{KL}(\cdot \parallel \cdot)$ seja chamado de "distância" em alguns trabalhos, não é uma distância entre distribuições no sentido conhecido do termo, porque não é simétrica e não satisfaz à desigualdade triangular. Segundo Kailath (1967), a divergência $D_{KL}(\cdot \parallel \cdot)$ mede a distância entre distribuições de probabilidade, sendo sensível às diferenças em suas formas. Ele também apontou que a divergência de KL é uma medida assimétrica que reflete a quantidade de informação perdida quando uma distribuição de probabilidade é aproximada pela outra.

Neste trabalho de tese, consideraremos uma simetrização da referida medida. Sejam dois vetores aleatórios, com o mesmo suporte, a distância Kullback-Leibler pode ser definida como se segue:

$$\begin{aligned} d_{KL}(\mathbf{x}, \mathbf{y}) &= \frac{1}{2} [D_{KL}(\mathbf{x} \parallel \mathbf{y}) + D_{KL}(\mathbf{y} \parallel \mathbf{x})] \\ &= \frac{1}{2} \int_{\mathcal{X}} [f_x(\mathbf{z}; \boldsymbol{\theta}_x) - f_y(\mathbf{z}; \boldsymbol{\theta}_y)] \log \frac{f_x(\mathbf{z}; \boldsymbol{\theta}_x)}{f_y(\mathbf{z}; \boldsymbol{\theta}_y)} d\mathbf{z}, \end{aligned} \quad (2.2)$$

em que $f_x(\mathbf{z}; \boldsymbol{\theta}_x)$ representa a função densidade de probabilidade (fdp) de $\boldsymbol{\theta}_x$, $f_y(\mathbf{z}; \boldsymbol{\theta}_y)$ representa a fdp de $\boldsymbol{\theta}_y$ e $\mathcal{X}$ representa o suporte da fdp.

Outra divergência importante que amplia a divergência KL é a divergência de Rényi (ERVEN; HARREMOËS, 2010), definida como segue.

Sejam $\mathbf{x}$ e $\mathbf{y}$ dois vetores p aleatórios cujo suporte está definido sobre $\mathcal{Z} \subseteq \mathbb{R}^p$ e indexados por vetores de parâmetros $\boldsymbol{\theta}_1$ e $\boldsymbol{\theta}_2$, a divergência de Rényi é dada por

$$D_{R:r}^1(\mathbf{x}||\mathbf{y}) = D_{R:r}^1(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2) = \frac{1}{r-1} \log \int_{\mathcal{Z}} f_{\mathbf{x}}^r(\mathbf{z}) f_{\mathbf{x}}^{1-r}(\mathbf{z}) d\mathbf{z}, \quad (2.3)$$

em que $r > 0$ e $r \neq 1$. Se $r \neq 1$, então a divergência de Rényi se aproxima da divergência de KL. Liese e Vajda (1987) estenderam a versão anterior da divergência Rényi como (para $r \in \mathbb{R}$ e $r \neq \{0, 1\}$):

$$D_{R:r}(\mathbf{x}||\mathbf{y}) = D_{R:r}(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2) = \frac{1}{r(r-1)} \log \int_{\mathcal{Z}} f_{\mathbf{x}}^r(\mathbf{z}) f_{\mathbf{x}}^{1-r}(\mathbf{z}) d\mathbf{z}. \quad (2.4)$$

Pode-se observar que $D_{R:r}(\mathbf{x}||\mathbf{y})$ não é uma medida simétrica, então para resolver essa problemática LIMA, NASCIMENTO e AMARAL (2021) derivaram a seguinte medida:

$$d_{R:r}(\mathbf{x}||\mathbf{y}) = \frac{1}{r(r-1)} \log \left\{ \frac{\exp[r(r-1) D_{R:r}(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2)] + \exp[r(r-1) D_{R:r}(\boldsymbol{\theta}_2||\boldsymbol{\theta}_1)]}{2} \right\}. \quad (2.5)$$

Para investigar a distribuição assintótica das medidas: $d_{KL}(\mathbf{x}, \mathbf{y})$, $D_{KL}(\mathbf{x} || \mathbf{y})$, $D_{R:r}(\mathbf{x} || \mathbf{y})$, e $d_{R:r}(\mathbf{x} || \mathbf{y})$, Salicrú et al. (1994) propuseram uma definição em que as consideravam como elementos em uma classe de divergências chamada (h, ϕ) -divergência. Essa classe é uma extensão da classe ϕ -divergência proposta por Csiszár (1967) e Ali e Silvey (1966).

A seguir, apresentamos um tratamento estatístico para testes de hipóteses com base na divergência Kullback-Leibler e medidas de distância.

Sejam $\mathbf{x}$ e $\mathbf{y}$ dois vetores aleatórios previamente definidos. Assumindo que ambos têm um suporte comum $\mathcal{Z} \subset \mathbb{R}^p$, a (h, ϕ) -divergência entre $f_{\mathbf{x}}(\mathbf{z}; \boldsymbol{\theta}_x)$ e $f_{\mathbf{y}}(\mathbf{z}; \boldsymbol{\theta}_y)$ é definida como

$$D_{\phi}^h(\mathbf{x} || \mathbf{y}) = h \left(E_{\mathbf{y}} \left[\phi \left(\frac{f_{\mathbf{x}}(\mathbf{y}; \boldsymbol{\theta}_x)}{f_{\mathbf{y}}(\mathbf{y}; \boldsymbol{\theta}_y)} \right) \right] \right) = h \left(\int_{\mathcal{Z}} \phi \left(\frac{f_{\mathbf{x}}(\mathbf{z}; \boldsymbol{\theta}_x)}{f_{\mathbf{y}}(\mathbf{z}; \boldsymbol{\theta}_y)} \right) f_{\mathbf{y}}(\mathbf{z}; \boldsymbol{\theta}_y) d\mathbf{z} \right), \quad (2.6)$$

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 atribuídos valor zero. Por uma escolha criteriosa de funções $[h(x)$ e $\phi(x)]$ surgem medidas de divergências conhecidas:

$$[h(x), \phi(x)] = \begin{cases} [x, x \log(x)], & \text{para } D_{KL}, \\ [1/2 x, (x-1) \log(x)], & \text{para } d_{KL} \\ \left[\frac{1}{r(r-1)} \log(r(r-1)x + 1), \frac{x^r - r(x-1) - 1}{r(r-1)} \right], & \text{para } D_{R:r}, \\ \left[\frac{1}{r(r-1)} \log\left(\frac{r(r-1)}{2}x + 1\right), \frac{x^r + x^{1-r} - x - 1}{r(r-1)} \right], & \text{para } d_{R:r}. \end{cases} \quad (2.7)$$

Ao considerar a definição de distâncias em termos das funções h e ϕ combinadas com os resultados propostos por Salicrú et al. (1994) na convergência em distribuição de uma estatística em função da classe de divergências (h, ϕ) para a distribuição Qui-quadrado, o seguinte lema pode ser demonstrado.

Lema 2.1.1 *Assuma como satisfeitas as condições de regularidade propostas em Salicrú et al. (1994). Se $\frac{m}{m+n} \xrightarrow{m,n \rightarrow \infty} \lambda \in (0, 1)$ e $\theta_x = \theta_y$, então*

$$\frac{2mn}{m+n} \frac{d_{\phi}^h(\widehat{\theta}_x, \widehat{\theta}_y)}{h'(0)\phi''(1)} \xrightarrow{m,n \rightarrow \infty} \chi_M^2, \quad (2.8)$$

em que $h'(0)$ é a derivada de 1ª ordem, quando $h = 0$, $\phi''(1)$ é a derivada de 2ª ordem, quando $\phi = 1$, “ $\xrightarrow{\mathcal{D}}$ ” denota a convergência em distribuição, M é a dimensão de θ_i , m é o primeiro tamanho amostral, n é o segundo tamanho amostral, $\widehat{\theta}_x$ é a EMV para θ_x baseado na primeira amostra, $\widehat{\theta}_y$ é a EMV para θ_y baseado na segunda amostra e “ $\sim \chi_M^2$ ” denota uma variável com distribuição qui-quadrado com M graus de liberdade.

Baseado no Lema 2.1.1, testes estatísticos para a hipótese nula $\theta_x = \theta_y$ podem ser derivados. Em particular, a seguinte estatística é considerada:

$$S_{\phi}^h(\widehat{\theta}_x, \widehat{\theta}_y) = \frac{2mnv}{m+n} d_{\phi}^h(\widehat{\theta}_x, \widehat{\theta}_y), \quad (2.9)$$

em que $v = 1/(h'(0)\phi''(1))$ é uma constante que depende das derivadas das funções h e ψ .

Para medidas D_{KL} , d_{KL} , $D_{R:r}$ e $d_{R:r}$, v assume valores 1/2, 1, 1 e 1, respectivamente. Sendo descrito dessa forma, um teste de hipótese geral pode ser descrito na seguinte proposição.

Proposição 2.1.2 *Sejam m e n tamanhos de amostras suficientemente grandes e $S_{\phi}^h(\widehat{\theta}_x, \widehat{\theta}_y) = s$, então a hipótese nula $\theta_x = \theta_y$ pode ser rejeitada a um nível η se $\Pr(\chi_p^2 > s) \leq \eta$, em que p é a estatística de teste calculada a partir de amostras observadas e η é o nível nominal pré-especificado.*

2.2 ATRIBUTO POLSAR: SPAN

Os sistemas PolSAR registram a amplitude e a fase dos sinais retrodifusos a partir de possíveis combinações de recepção linear e transmitem as seguintes polarizações: HH, HV, VH e VV (H para direção horizontal e V para direção vertical). O resultado é a matriz de espalhamento complexa dos alvos (scattering matrix complex), definida como:

$$\mathbf{S} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix}, \quad (2.10)$$

em que $S_{AB} \in \mathbb{C}$ significa o retorno devido à transmissão A e recepção B .

De acordo com Lee et al. (1994), os componentes co-polarizados S_{HH} e S_{VV} estão correlacionados, enquanto os componentes de polarização cruzadas S_{VH} e S_{HV} têm níveis iguais entre si.

Nesta seção, mediremos o impacto das correlações elevadas de distribuições bivariadas através de dois canais polarimétricos, assim descrevendo as características dos retornos do PolSAR. Então, seja $\mathbf{s} \in \{[S_{HH}, S_{VV}]', [S_{HH}, S_{HV}]', [S_{VV}, S_{HV}]'\} \subset \mathbb{C}^2$ um vetor de canais complexos de polarizações, em que $(\cdot)'$ indica o vetor transposto. Então, a matriz de covariância polarimétrica L -look $\mathbf{C}$ pode ser definida como

$$\mathbf{C} := \frac{1}{L} \sum_{i=1}^L \mathbf{s}_i \mathbf{s}_i^*, \quad (2.11)$$

em que $\mathbf{s}_i$ é o i -ésimo vetor de amostra single-look e $(\cdot)^*$ é o operador do conjugado transposto. Segundo Lee e Pottier (2009) a matriz $\mathbf{C}$ tem natureza hermitiana positiva definida, sendo assim é possível fazer decomposição espectral de $\mathbf{C}$ como:

$$\mathbf{C} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^* + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^*, \quad (2.12)$$

em que λ_i são os i -ésimos autovalores reais de $\mathbf{C}$, $\mathbf{e}_i$ são os autovetores ortonormais correspondentes e $\mathbf{e}_i^*$ são os autovetores conjugados transpostos correspondentes.

O poder de dispersão total da imagem (SPAN) pode ser definido como o poder de dispersão total dos dois mecanismos de dispersão em (2.12), ou seja,

$$\text{SPAN} := \lambda_1 + \lambda_2 = C_{11} + C_{22}, \quad (2.13)$$

em que C_{ij} é a (i, j) entrada de $\mathbf{C}$.

2.3 DETECTORES DE MUDANÇAS POLSAR

Nessa tese serão construídos três tipos de detectores de mudanças, dois detectores baseados em testes de hipóteses por divergências em probabilidade e um detector baseado na estatística da razão entre verossimilhanças, são eles:

- CD_{KL} : detector de mudança baseado na distância Kullback-Leibler;
- CD_R^α : detector de mudança baseado na distância Rényi de ordem α ;
- CD_{RV} : detector de mudança baseado na estatística da razão entre verossimilhanças.

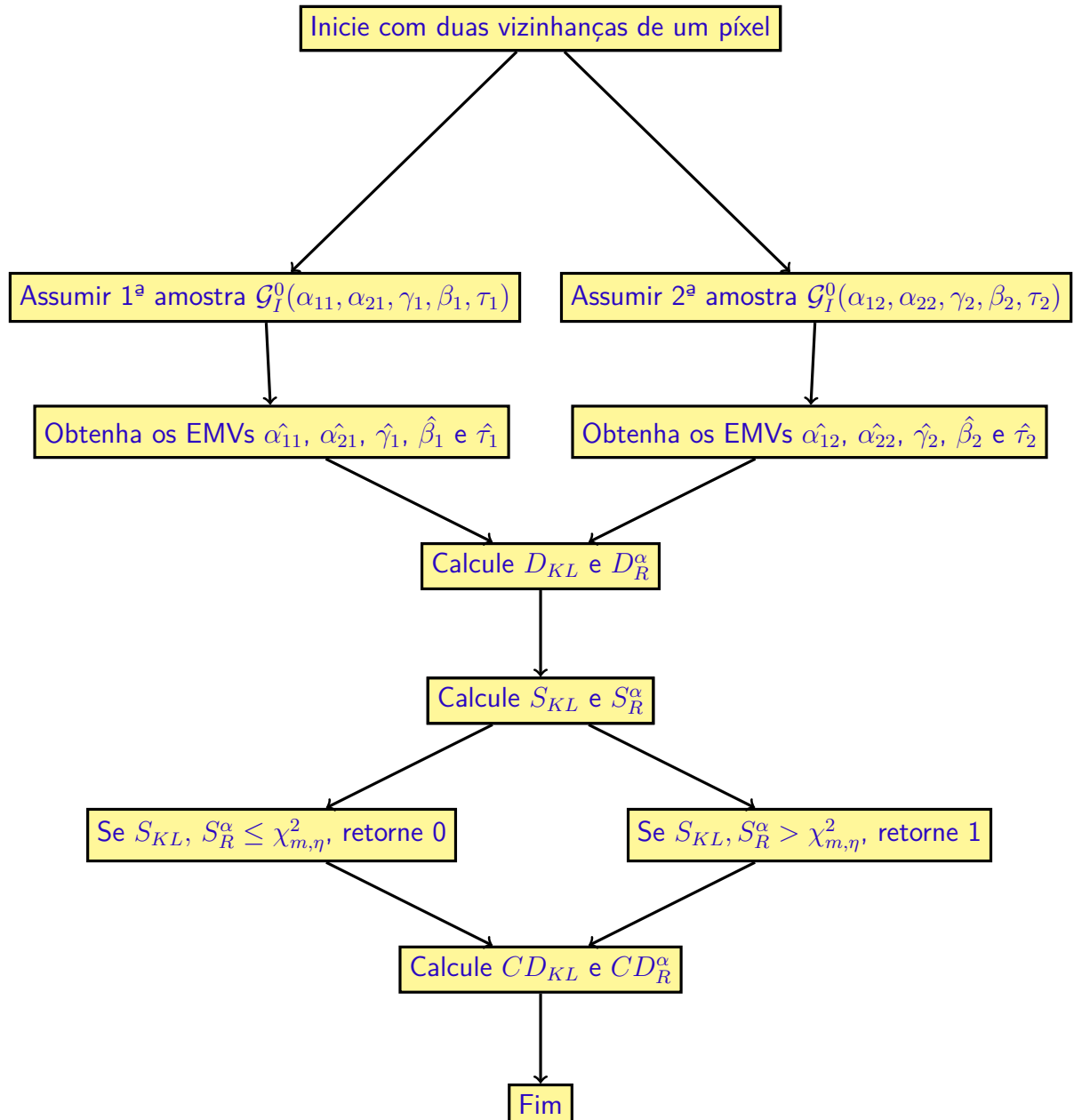
Esses detectores são algoritmos de processamento de imagem que conseguem identificar alterações entre duas imagens SAR, que podem ser causadas por mudanças na superfície da Terra, como em áreas urbanas densas. Esse tipo de detector compara as diferenças de intensidade entre as imagens procurando áreas onde houve uma mudança significativa na intensidade do sinal SAR entre as duas imagens. Para confirmar a discussão quantitativas dos mapas preditivos, usamos cinco métricas para avaliar o desempenho desses detectores:

1. *Falso Positivo (FP)*: Representa o número de pixels pelo mapa de referência, em que o algoritmo detectou mudanças entre as cenas da imagem, mas o número de pixels foi classificado erroneamente como sem mudança. Em estatística, é a probabilidade de um teste indicar erroneamente que um evento ocorreu, quando, na verdade não ocorreu.
2. *Falso Negativo (FN)*: Representa o número de pixels pelo mapa de referência, em que o algoritmo não detectou mudanças entre as cenas da imagem, mas o número de pixels foi classificado como alterado. Em estatística, é a probabilidade de um teste indicar erroneamente que um evento não ocorreu, quando na verdade ocorreu.
3. *Taxa de Falso Alarme*: $(FP+FN)/m$, onde m é o número de pixels inalterados conforme o detector.
4. *Taxa de Detecção*: TP/CG , onde TP é o número de pixels indicados como alteração tanto pelo RM como pelo detector, e CG é o número de pixels alterados conforme o detector.
5. *Coeficiente Kappa*: $\kappa = (A - B)/(1 - B)$, em que $A = 1 - p_{FP} - p_{FN}$ e $B = (p_{TP} + p_{FP})(p_{TP} + p_{FN}) + (p_{TN} + p_{FP})(p_{TN} + p_{FN})$, em que p_{FP} , p_{FN} , p_{TP} e

p_{TN} representa as proporções de falso positivo, falso negativo, verdadeiro positivo e verdadeiro negativo respectivamente, p_C é a proporção de pixels, sob a condição C em relação ao número total de pixels e TN é o número de pixels indicados como não havendo alteração tanto pelo RM como pelo detector.

A figura 2 mostra um diagrama para ilustrar como os testes baseados nas divergências de Kullback-Leibler e Rényi funcionam na prática até a tomada de decisão supondo o modelo da nova distribuição bivariada $\mathcal{G}_I^0 \text{MB}$.

Figura 2 – Diagrama de blocos para testes de hipóteses e detectores de mudança baseados nas divergências de Kullback-Leibler e Rényi. $\chi_{m,\eta}^2$, em que η representa quantil.



Fonte: Elaborada pelo autor (2023)

3 DISTRIBUIÇÃO BIVARIADA MCKAY APLICADA ÀS CARACTERÍSTICAS SAR BASEADAS EM SPAN E MODELAGEM MULTIPLICATIVA

Entre os métodos de modelagem de dados PolSAR, a modelagem multiplicativa (MM) tem grande importância devido ao seu envolvimento na formação física das imagens produzidas por tais sistemas. A MM diz que um dado PolSAR multilook $\mathbf{Z} \in \Omega := \{\mathbf{H} \in \mathbb{C}^p \times \mathbb{C}^p : \mathbf{H} = \mathbf{H}^* \mathbf{e} \mathbf{H} \succ \mathbf{0}\}$ vem do produto entre duas variáveis aleatórias independentes: a primeira aborda o ruído speckle multidimensional, digamos $\mathbf{Y} \in \Omega$, a segunda descreve o comportamento do terreno, digamos $X \in \mathbb{R}_+$, e o retorno é definido por $\mathbf{Z} = \mathbf{Y} \times X$, onde p é o número de canais de polarização. Vários modelos foram propostos dentro do MM, sendo a lei $\mathcal{G}_p^0(\mathcal{G}_p^0(\alpha, \Sigma, L))$ uma das distribuições mais importantes para $\mathbf{Z}$. Ela surge da hipótese de que $\mathbf{Y}$ segue a Lei de Wishart Complexa ($\mathcal{W}(\Sigma, L)$) e X a distribuição Gama Inversa ($\Gamma^{-1}(\alpha, (-1 - \alpha))$). Se $\mathbf{Z} \sim \mathcal{G}_p^0(\alpha, \Sigma, L)$, a sua densidade é dada por (FREITAS; FRERY; CORREIA, 2005)

$$f_{\mathbf{Z}}(\dot{\mathbf{Z}}; \alpha, \Sigma, L) = \frac{L^p \Gamma(Lp - \alpha)}{|\Sigma|^L (-\alpha - 1)^\alpha \Gamma(-\alpha) \Gamma_p(L)} |\dot{\mathbf{Z}}|^{L-p} [(-\alpha - 1) + L \text{tr}(\Sigma^{-1} \dot{\mathbf{Z}})]^{\alpha-Lp}, \quad (3.1)$$

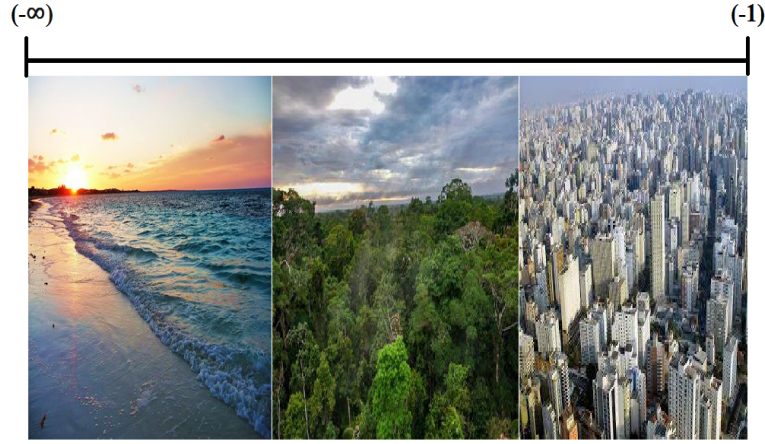
em que $\dot{\mathbf{Z}} \in \Omega$, $\alpha < -1$, L é o número de looks, Σ é a matriz hermitiana, α é o parâmetro de rugosidade, tr e $|\cdot|$ são o traço e o determinante da matriz, $\Gamma(\cdot)$ é a função gama e $\Gamma_p(\cdot)$ é a sua versão para a dimensão p ; portanto, Freitas, Frery e Correia (2005) determina L fisicamente como um modo de quantificar a razão sinal ruído e α o grau de homogeneidade do alvo. A figura 3 ilustra a associação do parâmetro de rugosidade α com o tipo de terreno.

A construção de algoritmos baseado em MM é de grande importância para detecção e classificação dos sinais em imagens PolSAR. Portanto, Lee et al. (1994) derivaram uma nova fdp conjunta para razão de intensidade e amplitude multilooks em relação aos dois elementos da diagonal S (ver seção 2.2), isto é, modelaram a intensidade e amplitude dos canais polarizados HH e VV. A fdp é definida como:

$$f_{R_1, R_2}(R_1, R_2) = \frac{L^{L+1} (R_1, R_2)^{\frac{L-1}{2}} \exp\left(-\frac{L(R_1+R_2)}{1-\rho^2}\right)}{\Gamma(L) (1-\rho^2) \rho^{L-1}} I_{L-1}\left(2L\sqrt{R_1 R_2} \frac{\rho}{1-\rho^2}\right), \quad (3.2)$$

em que $R_1, R_2, L > 0$ e $0 < \rho < 1$. Sendo ρ é o coeficiente de correlação, L é o número de looks, $I(\cdot)$ é a função de Bessel modificada do tipo I de ordem $n - 1$, R_i , para $i = 1, 2$ representa a intensidade normalizada multilook do i -ésimo canal polarimétrico dada pela expressão:

Figura 3 – Ilustração de alvos associados com valores diferentes do parâmetro de rugosidade α de uma distribuição $\mathcal{G}_m^0$.



Fonte: Elaborada pelo autor (2023)

$$R_i = \frac{1}{n} \sum_{k=1}^n \frac{|S_i(k)|^2}{\mathbb{E}(|S_i(k)|^2)}, \quad (3.3)$$

tal que S_i , para $i = 1, 2$ indica um dos elementos da matriz complexa S e $\mathbb{E}(\cdot)$ é o operador do valor esperado.

A função densidade de probabilidade para o par de intensidades normalizadas proposta por Lee et al. (1994) é definida como:

$$f_{B_1, B_2}(B_1, B_2) = \frac{(B_1, B_2)^{\frac{L-1}{2}} \exp\left(-\frac{(B_1+B_2)}{1-\rho^2}\right)}{\Gamma(L) (1-\rho^2) \rho^{L-1}} I_{L-1}\left(2\sqrt{B_1 B_2} \frac{\rho}{1-\rho^2}\right), \quad (3.4)$$

em que $B_1, B_2, L > 0$ e $0 < \rho < 1$, sendo $B_1 = LR_1$ e $B_2 = LR_2$.

Gao (2019) propôs um detector de mudanças para o produto de amplitudes multilook de dois canais polarimétricos baseados no SPAN, que segue uma distribuição de intensidade $\mathcal{G}^0$ aplicado à modelagem do fundo do mar. Ele também derivou uma nova fdp univariada para o produto normalizado $(B_1 B_2)$ de amplitudes multilook de dois canais polarimétricos. Essa distribuição foi proposta para melhorar a amplitude do retroespalhamento (ou a intensidade do alvo do radar) quando comparado com a desordem (clutter) do radar. A fdp é dada como:

$$f_{\varsigma}(\varsigma) = \frac{4\varsigma^L L^{L+1}}{\Gamma(L) (1-\rho^2) \rho^{L-1}} I_{L-1}\left(\frac{2L\rho}{1-\rho^2}\varsigma\right) K_0\left(\frac{2L}{1-\rho^2}\varsigma\right), \quad (3.5)$$

em que $\varsigma, L > 0, 0 < \rho < 1, \varsigma = \frac{\sqrt{B_1 B_2}}{L}$ e $K_0(\cdot)$ indica a função de Bessel modificada do tipo 2 de ordem 0.

Recorrendo às fórmulas assintóticas das funções de Bessel modificada Gao (2019) notou que a fdp (3.5) colapsa na da Gama definida como:

$$f_{\xi}(\xi) = \frac{\beta^L}{\Gamma(L)} \xi^{L-1} \exp(-\beta\xi), \quad (3.6)$$

em que $\xi, \beta, L > 0$, sendo $\beta = \frac{2L}{(1+\rho)\sigma_1\sigma_2}$ e $\sigma_i = \mathbb{E}(|S_i(k)|^2)$, para $i = 1, 2$.

A lei Gama e suas generalizações se destacam na literatura para modelar texturas (HACHICHA; CHAABANE, 2009), amplitude e intensidade (LI et al., 2011; HORTA et al., 2012) em imagens SAR capturadas de áreas de oceanos, urbanas e florestais densas. De acordo com Hagedorn et al. (2006), a distribuição gama multivariada é a lei natural para $C_{i,i}$ (C_{11} , SPAN) em cenários homogêneos. Uma das mais antigas distribuições gama bivariadas é a versão de McKay (NADARAJAH; GUPTA, 2006; GUPTA; NADARAJAH, 2006). Ela pode ser entendida como um desses dois resultados:

- (i) $\mathbf{x} = (X_1, X_2) := (V, U \times V)$ tais que $V \sim \Gamma(\alpha_1 + \alpha_2, 1/\gamma)$ e $U \sim \text{Beta}(\alpha_1, \alpha_2)$ são independentes e $U \times V \sim \Gamma(\alpha_1, 1/\gamma)$, em que $\mathbf{x}$ é o vetor de variável aleatória, “ $\sim \Gamma(\cdot, \cdot)$ ” denota a distribuição Gama e “ $\sim \text{Beta}(\cdot, \cdot)$ ” é a esperança da lei Beta, $X_1 = V$ e $X_2 = U \times V$ ou
- (ii) $\mathbf{x} = (X_3, X_4) := (V_0, V_0 + V_1)$ tais que $V_0 \sim \Gamma(\alpha_1, 1/\gamma)$ e $V_1 \sim \Gamma(\alpha_2, 1/\gamma)$ são independentes, em que $X_3 = V_0$ e $X_4 = V_0 + V_1$.

Em resumo, as suposições (i) e (ii) são ferramentas poderosas para modelar sistemas complexos em diferentes campos da ciência e da engenharia, e sua aplicação não se limita ao contexto do PolSAR. A suposição (i) é comum em muitos modelos matemáticos que descrevem o comportamento de sistemas naturais e artificiais. Por exemplo, ela é usada para modelar a distribuição de chuva em uma área geográfica específica na hidrologia (YUE; OUARDA; BOBÉE, 2001). A suposição (ii) também é usada em vários campos, como na teoria da confiabilidade, onde é usada para modelar a falha de sistemas e prever sua probabilidade de falha em um determinado tempo (NADARAJAH, 2005).

A Distribuição Gama Bivariada McKay (MBG) com parâmetro de localização $\alpha_1, \alpha_2 > 0$ e parâmetro de escala $\gamma > 0$ tem fdp dada por:

$$f(x_1, x_2; \alpha_1, \alpha_2, \gamma) = \frac{1}{\gamma^{\alpha_1+\alpha_2} \Gamma(\alpha_1) \Gamma(\alpha_2)} x_1^{\alpha_1-1} (x_2 - x_1)^{\alpha_2-1} \exp\left(-\frac{x_2}{\gamma}\right), \quad (3.7)$$

para $0 < x_1 < x_2$. Aqui, $\mathbb{E}(X_1) = \gamma \alpha_1$, $\mathbb{E}(X_2) = \gamma (\alpha_1 + \alpha_2)$, $\mathbb{V}\text{ar}(X_1) = \alpha_1 \gamma^2$, $\mathbb{V}\text{ar}(X_2) = (\alpha_1 + \alpha_2) \gamma^2$ e $\text{Cor}(X_1, X_2) = \alpha_2 / (\alpha_2 + \alpha_1)$, em que $\mathbb{E}(\cdot)$ é a esperança, $\mathbb{V}\text{ar}(\cdot)$ é a variância, e $\text{Cor}(\cdot)$ é a correlação.

Considere uma discussão sobre a obtenção das Estimativas por Máxima Verossimilhanças (MLE) para $\boldsymbol{\theta} = (\gamma, \alpha_1, \alpha_2)$. Seja $\mathbf{X} = [(x_{11}, x_{21})^\top, \dots, (x_{1n}, x_{2n})^\top]^\top$ uma amostra observada a partir de $(X_1, X_2) \sim \text{MBG}(\alpha_1, \alpha_2, \gamma)$, a log-verossimilhança associada $\boldsymbol{\theta}$ é

$$\begin{aligned} \ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}, (X_1, X_2)) &= -n(\alpha_1 + \alpha_2) \log \gamma - n \log \Gamma(\alpha_1) - n \log \Gamma(\alpha_2) \\ &+ (\alpha_1 - 1) \sum_{i=1}^n \log x_{1i} + (\alpha_2 - 1) \sum_{i=1}^n \log(x_{2i} - x_{1i}) - \frac{\sum_{i=1}^n x_{2i}}{\gamma}, \end{aligned} \quad (3.8)$$

em que n é o tamanho da amostra usado e $\log(\cdot)$ é a função logarítmica.

A partir de $[\partial \ell(\boldsymbol{\theta}) / \partial \gamma, \partial \ell(\boldsymbol{\theta}) / \partial \alpha_1, \partial \ell(\boldsymbol{\theta}) / \partial \alpha_2] \Big|_{\gamma=\hat{\gamma}, \alpha_1=\hat{\alpha}_1, \alpha_2=\hat{\alpha}_2} = [0, 0, 0]$, o MLE para γ é $\hat{\gamma} = \bar{x}_2 / (\hat{\alpha}_1 + \hat{\alpha}_2)$ de tal forma que $\bar{x}_2 = \sum_{i=1}^n x_{2i} / n$ e as EMVs para α_1 e α_2 são uma solução do sistema não-linear

$$\begin{cases} \frac{\sum_{i=1}^n \log x_{1i}}{n} = \log \hat{\gamma} + \Psi(\hat{\alpha}_1), \\ \frac{\sum_{i=1}^n \log(x_{2i} - x_{1i})}{n} = \log \hat{\gamma} + \Psi(\hat{\alpha}_2), \end{cases} \quad (3.9)$$

em que $\Psi(\cdot)$ é a função Psi.

No caso do sistema (3.9), que não possui uma solução analítica, podemos usar métodos de otimização não linear para encontrar os valores dos parâmetros que maximizam a verossimilhança do modelo. A partir deste ponto, utilizaremos o termo "EMVs" para nos referirmos ao estimador de máxima verossimilhança para o vetor de parâmetros $\boldsymbol{\theta}$, quando discutirmos teoria assintótica. Isso é comum em muitos contextos de modelagem estatística, pois o estimador de máxima verossimilhança é amplamente utilizado na prática e é frequentemente usado em teoria assintótica para mostrar propriedades estatísticas do estimador. Vamos continuar a usar essa notação ao longo da tese.

3.1 TESTE DA RAZÃO DE VEROSSIMILHANÇA PARA A DISTRIBUIÇÃO GAMA BIVARIADA

A estatística da razão da log-verossimilhança (RV) tem grande importância na inferência em modelos paramétricos. De acordo com Casella e Berger (2021), a estatística de teste baseada em $\mathcal{H}_0$ tem uma distribuição assintótica χ_q^2 , em que q é a diferença entre as dimensões dos espaços de parâmetros sob as hipóteses alternativas e nulas. Denotamos esses espaços por Θ_1 e Θ_0 , respectivamente.

A diferença entre as dimensões dos espaços de parâmetros Θ_0 e Θ_1 é uma medida da complexidade do modelo completo em relação à hipótese nula. A distribuição assintótica χ_q^2 da estatística de teste é obtida a partir da teoria assintótica da estimação de máxima verossimilhança, que mostra que a estatística de teste se aproxima de uma distribuição qui-quadrado com q graus de liberdade, quando o tamanho da amostra tende ao infinito. Isso permite que possamos realizar inferência estatística e testes de hipóteses sobre os parâmetros do modelo, mesmo que a distribuição real dos dados seja desconhecida ou difícil de obter analiticamente.

Sejam $\boldsymbol{\theta}_1 = (\alpha_{11}, \alpha_{21}, \gamma_1)$ e $\boldsymbol{\theta}_2 = (\alpha_{12}, \alpha_{22}, \gamma_2)$ vetores de parâmetros de dois pares aleatórios, digamos de duas amostras aleatórias de $MBG(\alpha_{11}, \alpha_{21}, \gamma_1)$ e $MBG(\alpha_{12}, \alpha_{22}, \gamma_2)$. Admita ainda que $[(X_{11}^{(1)}, X_{21}^{(1)}), \dots, (X_{1M1}^{(1)}, X_{2M1}^{(1)})]$ e $[(X_{12}^{(1)}, X_{22}^{(1)}), \dots, (X_{1M2}^{(1)}, X_{2M2}^{(1)})]$ são duas amostras aleatórias obtidas de $\boldsymbol{x}$ e $\boldsymbol{y}$. A estatística RV é definida como

$$S_{RV} = -2 \log \frac{\sup_{\boldsymbol{\theta} \in \Theta_0} \ell(\boldsymbol{\theta})}{\sup_{\boldsymbol{\theta} \in \Theta} \ell(\boldsymbol{\theta})}, \quad (3.10)$$

em que $\Theta = \Theta_0 \cup \Theta_1$ e $\Theta_0 \cap \Theta_1 = \emptyset$. Assim, temos,

$$\begin{aligned} \ell(\boldsymbol{\theta} \in \Theta_0) &= -(M1 + M2) \log(\omega_1) + (\alpha_1 - 1) \left[\sum_{i=1}^{M1} \log X_{1i}^{(1)} + \sum_{j=1}^{M2} \log X_{1j}^{(2)} \right] \\ &+ (\alpha_2 - 1) \left[\sum_{i=1}^{M1} \log(X_{2i}^{(1)} - X_{1i}^{(1)}) + \sum_{j=1}^{M2} \log(X_{2j}^{(2)} - X_{1j}^{(2)}) \right] \\ &- \frac{1}{\gamma_1} \left[\sum_{i=1}^{M1} X_{2i}^{(1)} + \sum_{j=1}^{M2} X_{2j}^{(2)} \right] \end{aligned}$$

e temos:

$$\begin{aligned}
\ell(\boldsymbol{\theta} \in \Theta) = & -M1 \log(\omega_1) - M2 \log(\omega_2) + (\alpha_{11} - 1) \sum_{i=1}^{M1} \log X_{1i}^{(1)} + (\alpha_{12} - 1) \sum_{j=1}^{M2} \log X_{1j}^{(2)} \\
& + (\alpha_{21} - 1) \sum_{i=1}^{M1} \log(X_{2i}^{(1)} - X_{1i}^{(1)}) + (\alpha_{22} - 1) \sum_{j=1}^{M2} \log(X_{2j}^{(2)} - X_{1j}^{(2)}) \\
& - \frac{1}{\gamma_1} \sum_{i=1}^{M1} X_{2i}^{(1)} - \frac{1}{\gamma_2} \sum_{j=1}^{M2} X_{2j}^{(2)}
\end{aligned}$$

Em que $\omega_i = \gamma_i^{\alpha_{1i} + \alpha_{2i}} \Gamma(\alpha_{1i}) \Gamma(\alpha_{2i})$ para $i = 1, 2$, α_{1i} para $i = 1, 2$ são os parâmetros de localização para a primeira distribuição considerada, α_{2i} para $i = 1, 2$ são os parâmetros de localização para a segunda distribuição considerada, γ_i para $i = 1, 2$ é o parâmetro de escala para a i -ésima distribuição.

3.2 NOVOS TESTES DE HIPÓTESE BASEADOS EM DIVERGÊNCIAS PARA A DISTRIBUIÇÃO GAMA BIVARIADA

Diversos trabalhos indicam que a divergência Kullback-Leibler é importante para quantificar contrastes nas imagens PolSAR (FRERY; NASCIMENTO; CINTRA, 2014; BOUHLEL; MÉRIC, 2020). Começamos por derivar essa divergência no teorema a seguir.

Teorema 3.2.1 *Sejam $\theta_1 = (\alpha_{11}, \alpha_{12}, \gamma_1)$ e $\theta_2 = (\alpha_{21}, \alpha_{22}, \gamma_2)$ vetores de parâmetros de dois pares aleatórios, digamos x e y . A divergência de Kullback-Leibler de (2.1) entre x e y para distribuição bivariada gama de McKay é*

$$\begin{aligned} D_{KL}(x \parallel y) = D_{KL}(\theta_1 \parallel \theta_2) = & \log \frac{\gamma_2^{\alpha_{21} + \alpha_{22}} \Gamma(\alpha_{21}) \Gamma(\alpha_{22})}{\gamma_1^{\alpha_{11} + \alpha_{12}} \Gamma(\alpha_{11}) \Gamma(\alpha_{12})} \\ & + \left\{ (\alpha_{11} - \alpha_{21}) [\Psi(\alpha_{11}) + \log \gamma_1] \right. \\ & \left. + (\alpha_{12} - \alpha_{22}) [\log \gamma_1 + \Psi(\alpha_{12})] - \left(\frac{\gamma_1 + \gamma_2}{\gamma_1 \gamma_2} \right) (\alpha_{11} + \alpha_{12}) \right\}, \end{aligned}$$

em que α_{1i} , $i = 1, 2$, são os parâmetros de localização para a primeira distribuição considerada, α_{2i} , $i = 1, 2$, são os parâmetros de localização para a segunda distribuição considerada, γ_i , $i = 1, 2$, é o parâmetro de escala para a i -ésima distribuição.

A prova está descrita no Apêndice A. No que se segue, apresentamos a distância KL associada.

Corolário 3.2.2 *Sejam $\theta_1 = (\alpha_{11}, \alpha_{12}, \gamma_1)$ e $\theta_2 = (\alpha_{21}, \alpha_{22}, \gamma_2)$ no contexto do Teorema 3.2.1, temos*

$$\begin{aligned} d_{KL}(\theta_1, \theta_2) &= \frac{1}{2} [D_{KL}(\theta_1 \parallel \theta_2) + D_{KL}(\theta_2 \parallel \theta_1)] \\ &= \frac{1}{2} \left\{ (\alpha_{11} - \alpha_{21}) [\Psi(\alpha_{11}) + \log \gamma_1 - \Psi(\alpha_{21}) - \log \gamma_2] \right. \\ &\quad + (\alpha_{12} - \alpha_{22}) [\Psi(\alpha_{12}) + \log \gamma_1 - \Psi(\alpha_{22}) - \log \gamma_2] \\ &\quad \left. - (\alpha_{11} + \alpha_{12} + \alpha_{21} + \alpha_{22}) \left[\frac{\gamma_2 + \gamma_1}{\gamma_1 \gamma_2} \right] \right\}. \end{aligned}$$

O teorema seguinte apresenta a expressão de divergência de Rényi para a distribuição MBG.

Teorema 3.2.3 *Sejam $\beta_1 = \alpha_{21} + r\alpha_{11} - r\alpha_{21}$, $\beta_2 = \alpha_{22} + r\alpha_{12} - r\alpha_{22}$ e $\eta = \beta_1 + \beta_2$. Sejam $\theta_1 = (\alpha_{11}, \alpha_{12}, \gamma_1)$ e $\theta_2 = (\alpha_{21}, \alpha_{22}, \gamma_2)$. A divergência de Rényi de (2.4) entre x e y para distribuição bivariada gama de McKay é*

$$D_{R:r}(\theta_1 \parallel \theta_2) = \frac{1}{r(1-r)} \left\{ \log \Gamma(\eta_1) + \log \Gamma(\eta_2) + (1-r) \left[\log \Gamma(\alpha_{21}) + \log \Gamma(\alpha_{22}) \right. \right. \\ \left. \left. - (\alpha_{21} + \alpha_{22}) [\log \gamma_1 - \log(\gamma_1 - \gamma_1 r + \gamma_2 r)] \right] \right. \\ \left. + r \left[(\alpha_{11} + \alpha_{12}) (\log \gamma_2 - \log(\gamma_1 - \gamma_1 r + \gamma_2 r) - \log \Gamma(\alpha_{11}) + \log \Gamma(\alpha_{12})) \right] \right\}.$$

A prova está descrita no Apêndice B. No que se segue, apresentamos a distância Rényi associada.

Corolário 3.2.4 *Sejam $\theta_1 = (\alpha_{11}, \alpha_{12}, \gamma_1)$ e $\theta_2 = (\alpha_{21}, \alpha_{22}, \gamma_2)$ no contexto do Teorema 3.2.3, temos*

$$d_{R:r}(\theta_1, \theta_2) = -\frac{\log 2}{r(r-1)} + \frac{1}{r(r-1)} \log \left\{ \Gamma(\beta_{21}) \Gamma(\beta_{22}) \right. \\ \times (1-r) \left[\Gamma(\alpha_{21}) \Gamma(\alpha_{22}) \gamma_1^{-(\alpha_{21}+\alpha_{22})} (\gamma_1 - \gamma_1 r + \gamma_2 r)^{-(\alpha_{21}+\alpha_{22})} \right] \\ \times r \left[\gamma_2^{\alpha_{11}+\alpha_{12}} (\gamma_1 - \gamma_1 r + \gamma_2 r)^{-(\alpha_{11}+\alpha_{12})} \Gamma(\alpha_{11})^{-1} \Gamma(\alpha_{12})^{-1} \right] \\ + \Gamma(\beta_{11}) \Gamma(\beta_{12}) r \left[\gamma_1^{\alpha_{21}+\alpha_{22}} (\gamma_2 - \gamma_2 r + \gamma_1 r)^{-(\alpha_{21}+\alpha_{22})} \Gamma(\alpha_{21})^{-1} \Gamma(\alpha_{22})^{-1} \right] \\ \left. \times (1-r) \left[\Gamma(\alpha_{11}) \Gamma(\alpha_{12}) \gamma_2^{-(\alpha_{11}+\alpha_{12})} (\gamma_2 - \gamma_2 r + \gamma_1 r)^{-(\alpha_{11}+\alpha_{12})} \right] \right\},$$

em que $\beta_{1i} = \alpha_{1i} + r\alpha_{1i} - r\alpha_{2i}$, $\beta_{2i} = \alpha_{2i} + r\alpha_{2i} - r\alpha_{1i}$ para $i = 1, 2$.

Seguem-se do Lema 2.1.1 e Proposição 2.1.2 (tomando $M = 3$), combinados com Teoremas 3.2.1 and 3.2.3 e os Corolários 3.2.2 e 3.2.4, dois novos testes de hipóteses para a distribuição MBG.

- Estatística de Kullback-Leibler - $S_{KL}(\cdot, \cdot)$:

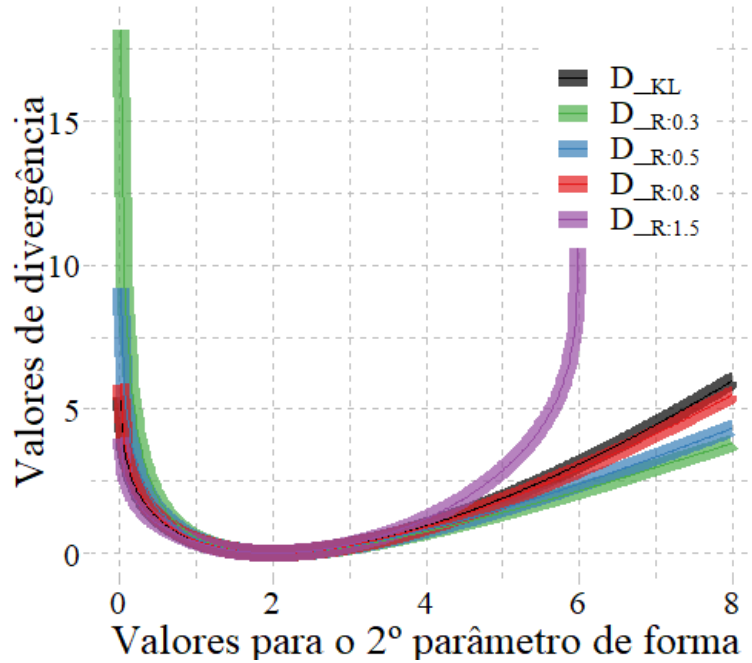
$$S_{KL}([\hat{\alpha}_{11}, \hat{\alpha}_{12}, \hat{\gamma}_1], [\hat{\alpha}_{21}, \hat{\alpha}_{22}, \hat{\gamma}_2]) \\ = \frac{2mn}{m+n} d_{KL}([\hat{\alpha}_{11}, \hat{\alpha}_{12}, \hat{\gamma}_1], [\hat{\alpha}_{21}, \hat{\alpha}_{22}, \hat{\gamma}_2]) \\ \xrightarrow[m, n \rightarrow \infty]{\mathcal{D}} \chi_3^2. \quad (3.11)$$

- Estatística de Rényi de ordem r - $S_R^r(\cdot, \cdot)$:

$$\begin{aligned}
 S_R^r([\hat{\alpha}_{11}, \hat{\alpha}_{12}, \hat{\gamma}_1], [\hat{\alpha}_{21}, \hat{\alpha}_{22}, \hat{\gamma}_2]) \\
 &= \frac{2mn}{m+n} d_R^r([\hat{\alpha}_{11}, \hat{\alpha}_{12}, \hat{\gamma}_1], [\hat{\alpha}_{21}, \hat{\alpha}_{22}, \hat{\gamma}_2]) \\
 &\xrightarrow[m, n \rightarrow \infty]{\mathcal{D}} \chi_3^2.
 \end{aligned} \tag{3.12}$$

A figura 4 mostra curvas da divergência $D_\bullet(\theta_1 || \theta_2(x))$ para a lei MBG, em que $\theta_1 = (1, 2, 1)$ e $\theta_2(x) = (1, x, 1)$ e x varia na faixa de $(0, 8)$. É perceptível que a divergência de Rényi converge para a divergência de KL quando r tende a um, como esperado. Pequenos valores de r tornam o aumento da divergência Rényi mais sensível para $x < 2$, enquanto, para valores altos de r , a divergência Rényi mais sensível para $x > 2$.

Figura 4 – Medidas de distância entre duas variáveis aleatórias MBG em função de α_1



Fonte: Elaborada pelo autor (2023)

Visando compreender melhor os resultados teóricos obtidos nessa seção, a seguir será realizado um estudo de simulação (ver Seção 3.3) que ajudará a fornecer uma compreensão mais profunda dos testes e de como ele deve ser aplicado. Teremos outra seção (ver Seção 3.4), cuja ênfase será na aplicação prática do algoritmo, mostrando como ele pode ser usado em situações reais, especificamente, para resolver problemas de detecção de mudanças em dados reais.

3.3 ESTUDO DE SIMULAÇÃO

Nesta seção, realiza-se um estudo de simulação para avaliar o comportamento dos testes propostos na Seção 3.2. Inicialmente, quantificamos o desempenho das ferramentas estatísticas propostas para construir detectores de mudança para as imagens PolSAR com base em teste de duas amostras em divergência Kullback-Leibler e Rényi (de ordem $\beta = 0.2, 0.5, 0.9$, e 1.5), e, estatística da razão entre verossimilhanças (RV).

Quanto à escolha dos valores de β , no contexto da classificação dos dados PolSAR, descobrimos que o método de Rényi funciona melhor quando a ordem está dentro do espaço $(0, 1)$ (RODRÍGUEZ et al., 2020; FERREIRA; COELHO; NASCIMENTO, 2021). Então, optamos por definir a maioria das ordens de Rényi entre $(0, 1)$.

Os métodos KL, Rényi e RV foram executados com 1000 iterações de Monte Carlo e com tamanhos de amostra $N = 9, 25, 49, 81$ e 121 . Na prática, definir janelas grandes em imagens significa misturar áreas estatisticamente diferentes, o que pode impedir a detecção de mudanças.

No entanto, em termos de inferência estatística, a variação do tamanho da amostra é uma forma para testar como as propriedades assintóticas dos testes se comportam. Como figura de mérito, utilizamos os tamanhos de testes empíricos nos níveis nominais de 1%, 5%, e 10%, valores usados frequentemente em estudos estatísticos.

Com relação aos parâmetros em $H_0 : (\alpha_{11}, \alpha_{12}, \gamma_1) = (\alpha_{21}, \alpha_{22}, \gamma_2)$, eles são definidos usando EMVs obtidos do campo de ensaio agrícola de Foulum em Jutland (Dinamarca) e registrado pelo EMISAR. O EMISAR é um radar de abertura sintética de banda C e L totalmente polarimétrico (ou seja, 4 canais complexos por frequência), projetado para sensoriamento remoto com altas demandas de resolução (2 m), discriminação de polarização e calibração radiométrica e polarimétrica absoluta. Podemos observar as características de cada amostra na Tabela 1, que mostra os valores das estimativas de $(\alpha_1, \alpha_2, \gamma)$ e correlações associadas para cada sub-região nas figuras 6a, 6b, e 6c.

Observando as curvas de nível das figuras 6a, 6b, e 6c, percebe-se que a Cena 2 (*área florestal* discutido em Skriver et al. (2005)) em termos de magnitude da forma, tem uma textura mais pronunciada que a Cena 1 (*trigo de inverno* o que foi discutido em Wang, Han e Deng (2017)), que tem uma textura mais evidente do que na Cena 3 (*colheitas largas de folhas* discutido em Zhang et al. (2015)). Vale a pena notar que os valores de γ estão relacionados diretamente aos das correlações associadas.

Tabela 1 – Valor dos parâmetros para os cenários simulados.

	α_1	α_2	γ	Correlação $\left(\sqrt{\frac{\alpha_1}{\alpha_1 + \alpha_2}}\right)$
Cenário 1	6.443582280	14.915205474	0.006888375	54.92%
Cenário 2	15.61487646	6.24223856	0.01936444	84.52%
Cenário 3	2.762274772	8.125188024	0.001330168	50.36%

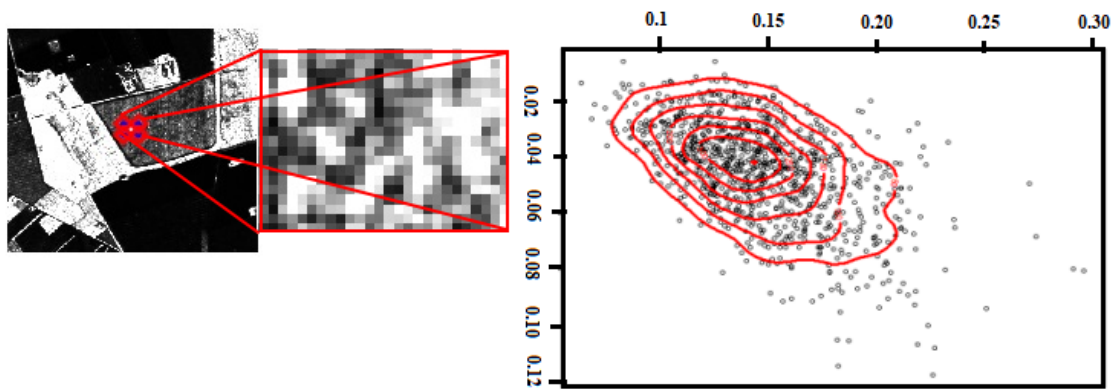
Fonte: Elaborada pelo autor (2023)

As taxas de rejeição para os níveis nominais podem ser vistas na Tabela 2. Nela, observam-se os tamanhos dos testes empíricos sob a hipótese nula, descrita em 2.1.2. Nesse estudo, os melhores resultados são destacados em cinza. Analisando globalmente, o teste baseado na estatística $S_R^{0,5}$ produziu os melhores resultados, pois seus valores foram os mais próximos dos níveis nominais assumidos. Os testes baseados em $S_R^{1,5}$ e S_{RV} não tiveram um bom desempenho.

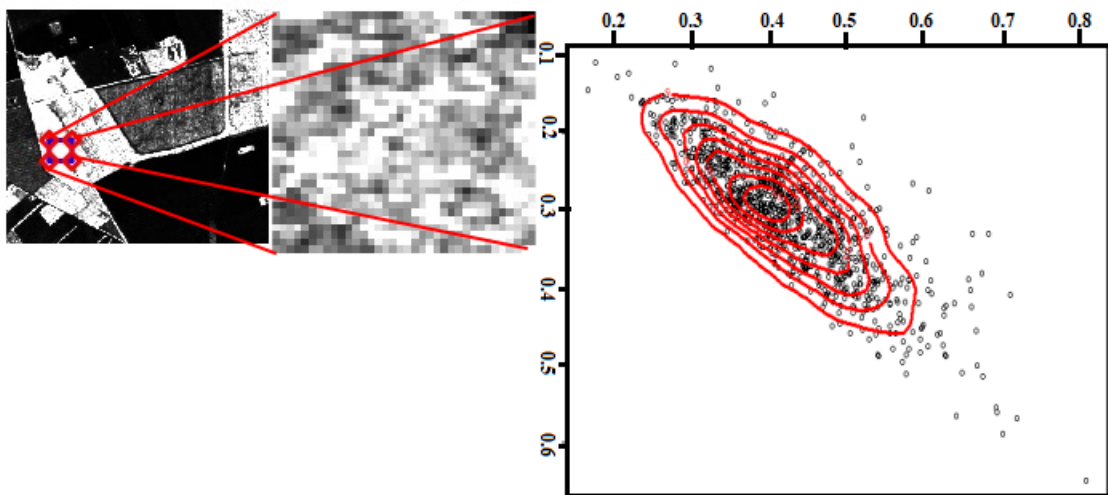
Analisando o desempenho dos testes propostos por região, podem-se destacar os seguintes testes:

- No nível nominal 10%, as estatísticas de Rényi $S_R^{0,2}$, $S_R^{0,5}$ e $S_R^{0,9}$ foram associados ao primeiro cenário; para o segundo cenário, foram as estatísticas S_{KL} , $S_R^{0,2}$ e $S_R^{0,9}$; para o terceiro cenário, apenas a estatística de Rényi $S_R^{0,5}$.
- Com respeito ao nível nominal de 5%, para a Cena 1, as estatísticas $S_R^{0,2}$ e $S_R^{0,5}$ apresentaram boas estimativas bem próximas aos níveis nominais; para a Cena 3, apenas a estatística $S_R^{0,5}$, e, por fim, na segunda Cena, apenas $S_R^{1,5}$ com tamanho $N = 81$. Além disso, o método S_{RV} não obteve uma boa estimativa.
- No nível 1%, as estatísticas de teste S_{KL} , $S_R^{0,2}$, $S_R^{0,5}$ e $S_R^{0,9}$ obtiveram as melhores estimativas para o primeiro e segundo cenário. No terceiro cenário, destaca-se apenas $S_R^{0,5}$.

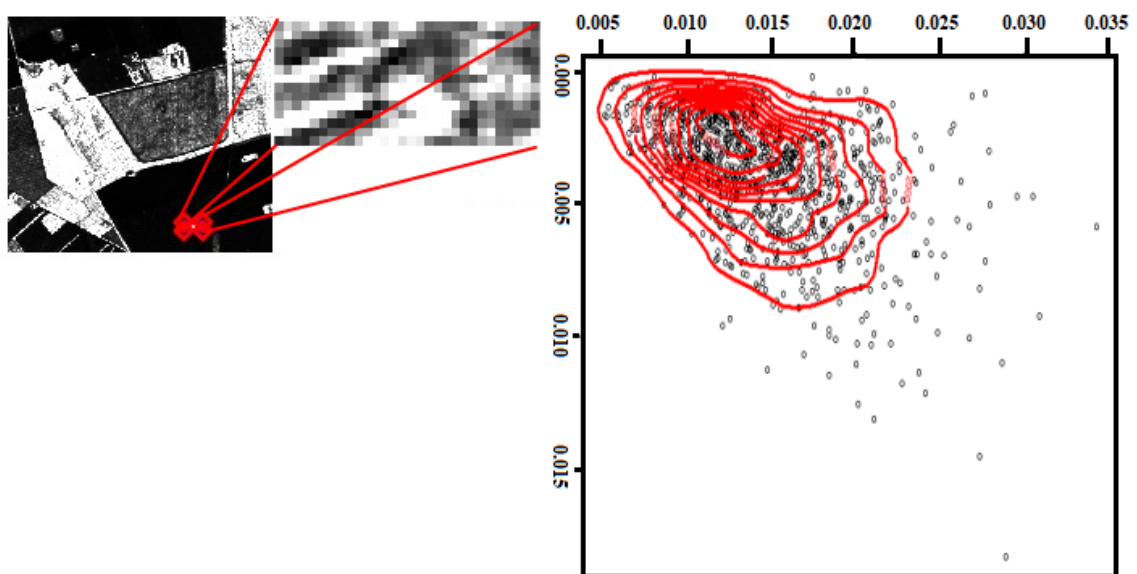
Figura 5 – Cenários de simulação obtida do campo de ensaio agrícola de Foulum em Jutland (Dinamarca), gráfico de dispersão de valores de pares (HH, HH+VV), e curvas de nível associadas.



(a) Cenário 1.



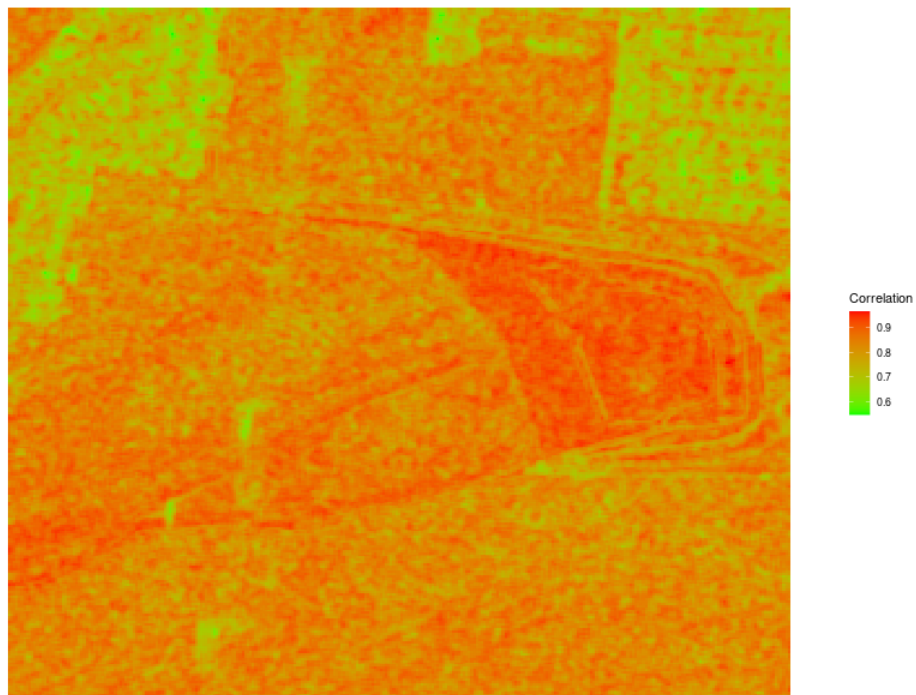
(b) Cenário 2.



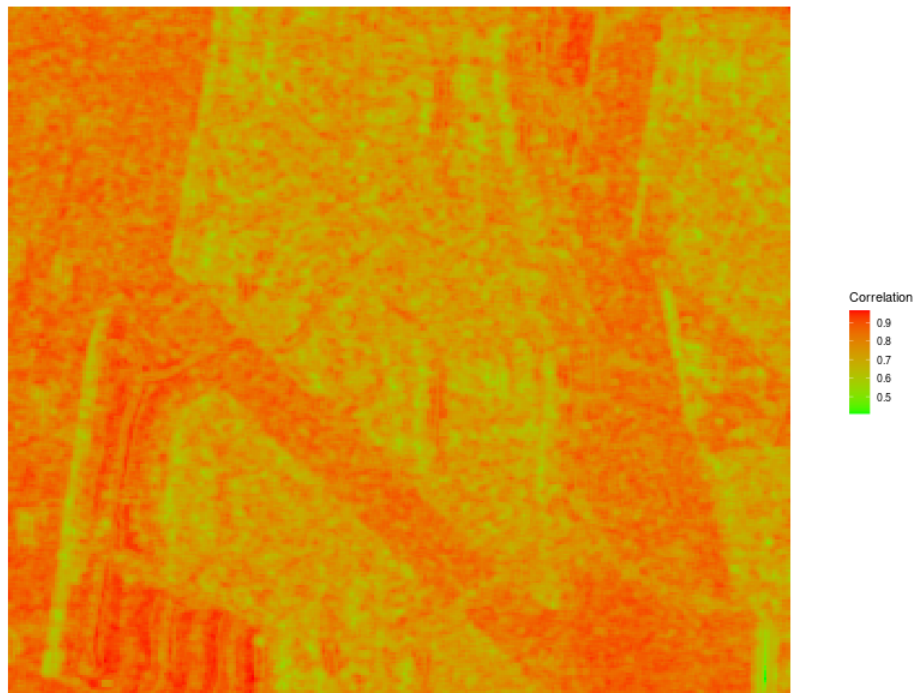
(c) Cenário 3

Fonte: Elaborada pelo autor (2023)

Figura 6 – Mapas de correlação MGB para as imagens de Los Angeles (USA), capturadas em 2015.



(a) Primeira região.



(b) Segunda região.

Fonte: Elaborada pelo autor (2023)

Tabela 2 – Taxas de Rejeição de Testes de Divergência sob $H0 : (\alpha_{11}, \alpha_{12}, \gamma_1) = (\alpha_{21}, \alpha_{22}, \gamma_2)$.

Cenários	N	Nível nominal 10%						Nível nominal 5%						Nível nominal 1%					
		KL	R0.2	R0.5	R0.9	R1.5	RV	KL	R0.2	R0.5	R0.9	R1.5	RV	KL	R0.2	R0.5	R0.9	R1.5	RV
1	9	0,172	0,160	0,149	0,165	0,240	0,188	0,113	0,095	0,089	0,103	0,240	0,144	0,045	0,033	0,025	0,039	0,240	0,087
	25	0,122	0,117	0,115	0,118	0,153	0,140	0,067	0,061	0,060	0,065	0,091	0,104	0,017	0,016	0,015	0,016	0,036	0,049
	49	0,110	0,109	0,108	0,110	0,119	0,130	0,055	0,054	0,054	0,055	0,065	0,100	0,009	0,009	0,009	0,009	0,012	0,046
	81	0,101	0,099	0,098	0,099	0,107	0,122	0,053	0,051	0,050	0,052	0,063	0,085	0,010	0,010	0,009	0,010	0,016	0,043
	121	0,101	0,100	0,100	0,100	0,106	0,136	0,052	0,049	0,048	0,051	0,059	0,091	0,011	0,011	0,010	0,011	0,013	0,038
2	9	0,169	0,156	0,151	0,164	0,246	0,196	0,110	0,089	0,082	0,095	0,246	0,154	0,039	0,029	0,022	0,033	0,246	0,083
	25	0,111	0,106	0,102	0,107	0,148	0,157	0,056	0,052	0,051	0,053	0,087	0,101	0,012	0,011	0,011	0,012	0,033	0,052
	49	0,114	0,113	0,111	0,114	0,124	0,146	0,067	0,065	0,062	0,065	0,077	0,104	0,020	0,018	0,018	0,018	0,023	0,043
	81	0,100	0,099	0,097	0,099	0,107	0,141	0,049	0,049	0,049	0,049	0,055	0,104	0,011	0,010	0,010	0,010	0,014	0,044
	121	0,114	0,112	0,111	0,113	0,118	0,125	0,060	0,060	0,059	0,060	0,071	0,084	0,013	0,013	0,013	0,013	0,015	0,035
3	9	0,078	0,063	0,060	0,068	0,228	0,184	0,048	0,041	0,038	0,041	0,228	0,144	0,014	0,011	0,008	0,012	0,228	0,069
	25	0,069	0,067	0,066	0,067	0,067	0,154	0,036	0,033	0,032	0,033	0,033	0,111	0,007	0,006	0,005	0,007	0,007	0,050
	49	0,104	0,101	0,100	0,102	0,125	0,132	0,051	0,051	0,050	0,051	0,070	0,085	0,014	0,014	0,013	0,014	0,018	0,032
	81	0,111	0,109	0,108	0,111	0,118	0,122	0,053	0,052	0,052	0,052	0,060	0,081	0,008	0,008	0,007	0,008	0,017	0,033
	121	0,110	0,109	0,108	0,109	0,118	0,117	0,053	0,051	0,051	0,052	0,057	0,077	0,014	0,014	0,014	0,014	0,018	0,030

Fonte: Elaborada pelo autor (2023)

3.4 APLICAÇÃO EM DADOS REAIS

Nesta seção, realiza-se uma aplicação a dados reais extraídos de imagens PolSAR obtidas do sensor de radar de abertura sintética para veículos aéreos (Uninhabited Aerial Vehicle Synthetic Aperture Radar - UAVSAR) da JPL. O nosso objetivo é resolver problemas de detecção de mudanças, utilizando as estatísticas propostas. O problema é descrito da seguinte forma: a partir de uma imagem PolSAR registradas em dois momentos diferentes, cria-se um mapa mostrando uma alteração no tipo de dados SAR. Portanto, serão utilizadas duas cenas capturadas pelo sensor UAVSAR referente às áreas urbanas densas da cidade de Los Angeles (CA, EUA), cujas mudanças são causadas pelo processo de urbanização.

O primeiro par - visto nas figuras 8a e 8b- foi capturada em 23 de abril de 2009 e 3 de maio de 2015, respectivamente. O segundo par mostrado nas figuras 10a e 10b- foi capturada em 23 de abril de 2009 e 3 de maio de 2015, respectivamente. A seguir, os detectores de mudança propostos são usados para detectar mudanças na região que ocorreram entre os dois períodos diferentes.

No entanto, aplicam-se os métodos propostos em imagens reais para verificar a eficiência dos testes empíricos em cenários complexos, como em áreas urbanas e florestas densas; que são diretamente afetadas pelo ruído speckle, além de facilitar o processo de amostragem à medida que o algoritmo tenta descrever claramente os dados da imagem no processo de detecção de mudanças. Para obter os resultados, utilizamos uma janela de tamanho 3×3 na vizinhança de pixels e os testes baseados em divergência sob o nível de significância 1×10^{-8} .

O janelamento de pixels é uma técnica utilizada em processamento de imagem para realçar ou enfatizar determinadas características, ou regiões de interesse em uma imagem. Já a vizinhança de pixels é uma região ao redor de um determinado pixel em uma imagem. Essa região pode ser definida como um quadrado ou um retângulo com um número ímpar de linhas e colunas. A vizinhança de um pixel é geralmente usada em algoritmos de processamento de imagem que envolvem filtragem e detecção de bordas de uma imagem. Nesse estudo, o tamanho da vizinhança de pixels e o nível nominal foram escolhidos para verificar a eficiência do método em detectar mudanças em terrenos mais densos, a exemplo de regiões urbanas cujo nível do ruído speckle é bastante elevado.

A figura 6 mostra os mapas de correlação MGB quantificados para as regiões de Los Angeles (CA, USA) em 3 de maio de 2015. As regiões com tons mais escuros (vermelho) representam áreas com altas correlações, enquanto as regiões destacadas com tons mais claros

(verde) possuem baixas correlações. O primeiro mapa pode representar uma análise em que as correlações são baixas, o que pode resultar em uma imagem mais ruidosa com regiões que não apresentam padrões claros de correlação. Já o segundo mapa pode representar uma análise em que as correlações são mais fortes, resultando em uma imagem mais clara e menos ruidosa.

As figuras 7-8 exibem mapas preditivos para a primeira aplicação. Os mapas preditivos mostram a distribuição espacial de áreas urbanas, áreas agrícolas, áreas florestais, entre outros. Eles são criados a partir de modelos estatísticos e algoritmos de classificação que processam os dados PolSAR e os convertem em mapas temáticos de alta resolução espacial. Especificamente nesta seção os mapas serão aplicados para verificar a detecção de mudanças do processo de urbanização da cidade de Los Angeles (CA-EUA).

Observa-se que os detectores baseados na distância de Rényi $CD_R^{0,2}$ e $CD_R^{0,9}$ produziram mapas, descritivos com representações próximas do mapa de referência para os pares de canais (HH-HV, HH-HV e HV-VV). Esses mapas estão destacados e podem ser vistos nas figuras 8g, 9d e 9g respectivamente. Os testes avaliativos com os resultados quantitativos estão descritos na Tabela 3.

O mapa de referência reflete o nível de mudanças entre duas imagens idênticas registradas em momentos distintos, ou seja, quantifica o processo de urbanização de uma imagem. Os resultados quantitativos acima podem ser confirmados observando o coeficiente Kappa para cada par de canais. Para os pares de canais HH-HV, o coeficiente Kappa, com valores próximos a 1, foi obtido por $CD_R^{0,9}$ com 65,20%, e o mesmo detector apresentou o melhor desempenho para o par HH-VV, com o coeficiente Kappa 62,29%. Para o par de canais HV-VV, o melhor detector foi $S_R^{0,2}$, com o coeficiente Kappa de 66,17%.

Com relação à segunda Cena da cidade de Los Angeles (CA, USA) os mapas podem ser vistos nas figuras 9-10. Observando o mapa de referência para essa Cena (ver figura 10c), nota-se que as maiores mudanças entre 2009 e 2015 estão em regiões localizadas no topo da imagem. No entanto, percebe-se que o detector baseado na divergência de Kullback Leibler CD_{KL} produziu mapas com excelente cobertura do processo de urbanização em Los Angeles (Califórnia -EUA). Os resultados quantitativos estão descritos na Tabela 4. O CD_{KL} produziu um coeficiente kappa de 63.57% para o par de canais HH-HV, 58.63% para o par de canais HH-VV e 64.01% para o par de canais HV-VV. Em resumo, os detectores produziram mapas preditivos com baixo ruído e que se assemelham aos mapas de referência, indicando uma boa captura do processo de urbanização da cidade de Los Angeles.

Tabela 3 – Desempenho dos detectores - Cena 1 - Los Angeles (CA, USA).

Detectors	FP (%)	FN (%)	FA (%)	DR (%)	κ (%)
HH-HV					
S_{KL}	5.248	4.317	58.936	94.847	63.991
$S_R^{0.2}$	2.895	5.699	68.781	93.488	64.147
$S_R^{0.5}$	2.342	6.299	76.195	92.895	62.704
$S_R^{0.9}$	3.675	5.019	62.303	94.167	65.201
S_{RV}	20.032	10.191	146.836	87.168	-12.342
HH-VV					
S_{KL}	4.956	4.908	64.281	94.202	61.985
$S_R^{0.2}$	2.614	6.584	81.195	92.575	60.287
$S_R^{0.5}$	2.101	7.329	93.642	91.851	57.691
$S_R^{0.9}$	3.405	5.753	70.715	93.391	62.292
S_{RV}	19.271	10.212	148.896	87.267	-12.151
HV-VV					
S_{KL}	6.988	3.342	54.531	95.877	63.685
$S_R^{0.2}$	4.351	4.404	57.425	94.804	66.173
$S_R^{0.5}$	3.692	4.909	61.088	94.286	65.691
$S_R^{0.9}$	5.262	3.909	55.081	95.311	65.848
S_{RV}	22.696	10.157	141.113	86.761	-13.203

Fonte: Elaborada pelo autor (2023)

Tabela 4 – Desempenho dos detectores - Cena 2 - Los Angeles (CA, USA).

Detectores	FP (%)	FN (%)	FA (%)	DR (%)	κ (%)
HH-HV					
S_{KL}	4.646	2.947	61.033	96.634	63.577
$S_R^{0.2}$	1.178	5.863	116.257	93.759	54.566
$S_R^{0.5}$	0.948	6.435	140.496	93.208	50.342
$S_R^{0.9}$	1.507	5.201	95.176	94.405	58.787
S_{RV}	2.202	4.209	73.389	95.389	62.98
HH-VV					
S_{KL}	4.733	3.576	69.833	95.941	58.632
$S_R^{0.2}$	1.159	6.566	144.817	93.064	48.258
$S_R^{0.5}$	0.947	7.154	178.676	92.506	43.353
$S_R^{0.9}$	1.473	5.883	116.176	93.719	53.184
S_{RV}	2.166	4.884	87.866	94.691	58.629
HV-VV					
S_{KL}	6.588	2.483	61.103	97.084	64.009
$S_R^{0.2}$	2.041	5.077	92.379	94.499	57.609
$S_R^{0.5}$	1.741	5.721	110.349	93.864	53.514
$S_R^{0.9}$	2.487	4.405	78.109	95.169	60.996
S_{RV}	3.351	3.507	64.779	96.078	59.499

Fonte: Elaborada pelo autor (2023)

Figura 7 – Mapas de detecção de mudanças do SPAN HH-HV para região de Los Angeles (CA, USA).



(a) Cena 1- 2009.

(b) Cena 1 - 2015.

(c) Mapa de referência.



(d) S_{KL} - HH-HV.



(e) $S_R^{0.2} - 3 \times 3$ - HH-HV.



(f) $S_R^{0.5}$ - HH-HV.



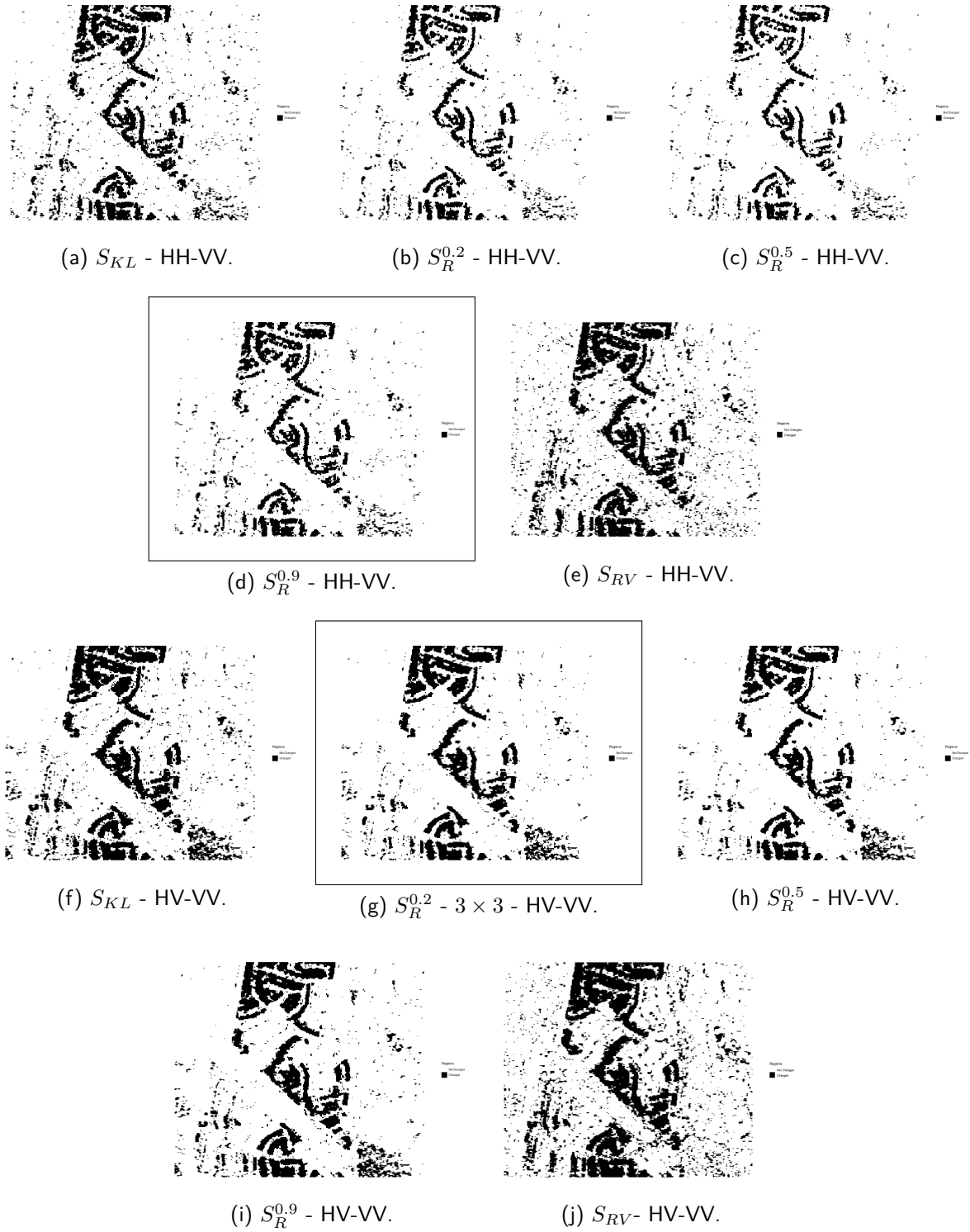
(g) $S_R^{0.9}$ - HH-HV.



(h) S_{RV} - HH-HV.

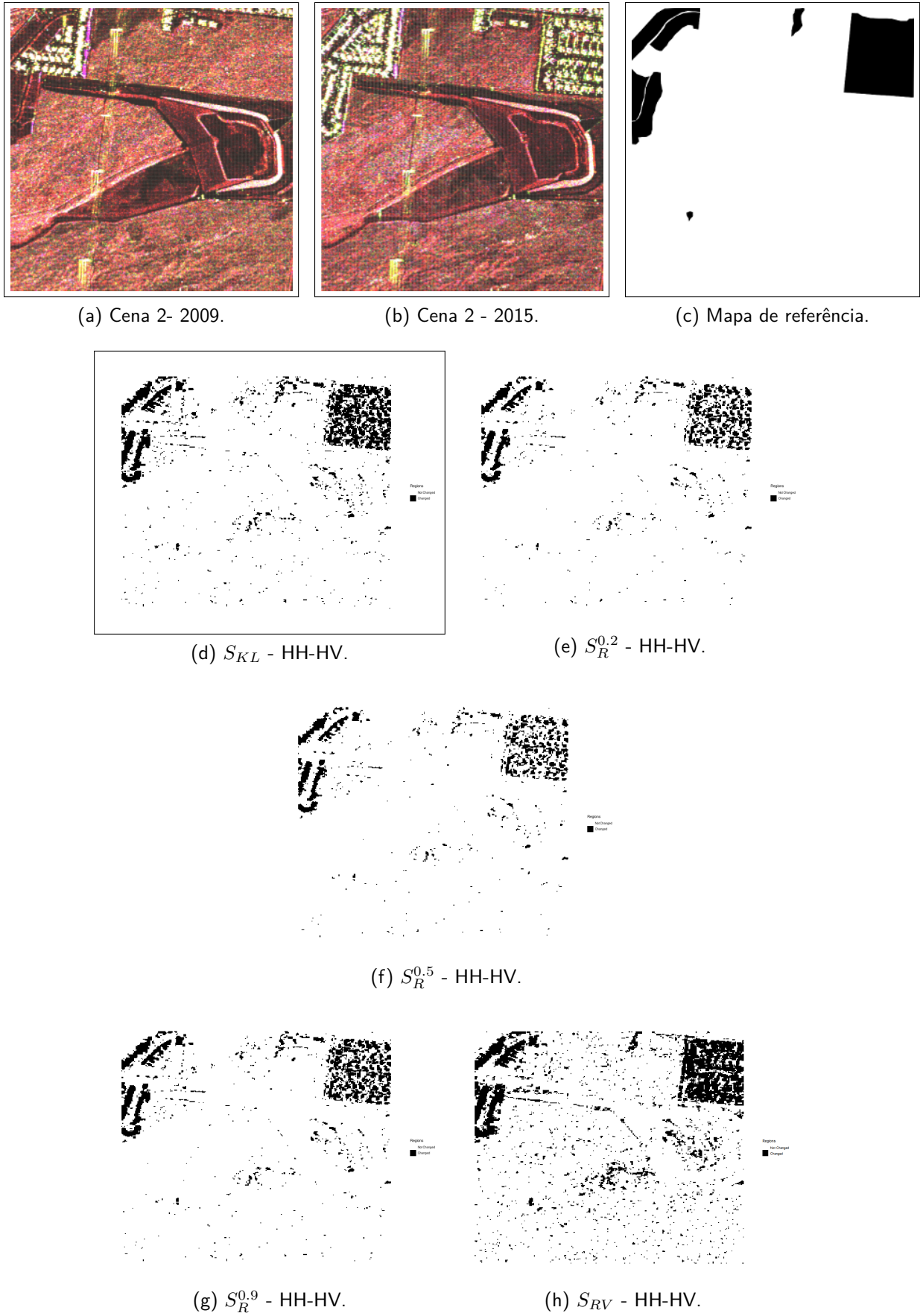
Fonte: Elaborada pelo autor (2023)

Figura 8 – Mapas de detecção de mudanças do SPAN HH-VV e HH-VV para região de Los Angeles (CA, USA).



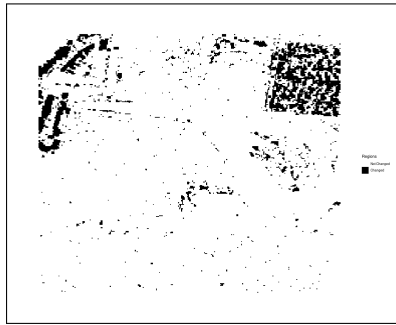
Fonte: Elaborada pelo autor (2023)

Figura 9 – Mapas de detecção de mudanças do SPAN HH-HV para região de Los Angeles (CA, USA).

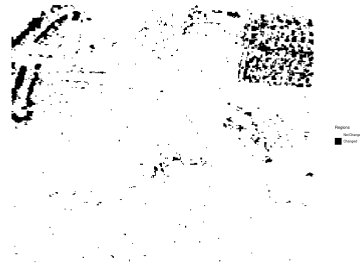


Fonte: Elaborada pelo autor (2023)

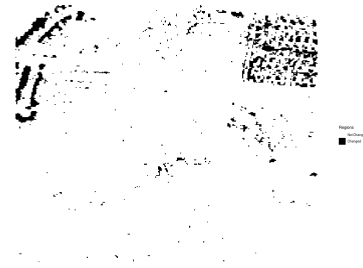
Figura 10 – Mapas de detecção de mudanças do SPAN HH-VV e HH-VV para região de Los Angeles (CA, USA).



(a) S_{KL} - HH-VV.



(b) $S_R^{0.2}$ - HH-VV.



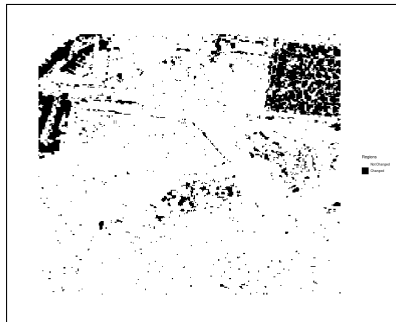
(c) $S_R^{0.5}$ - HH-VV.



(d) $S_R^{0.9}$ - HH-VV.



(e) S_{RV} - HH-VV.



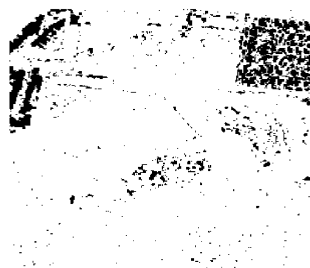
(f) S_{KL} - HV-VV.



(g) $S_R^{0.2}$ - HV-VV.



(h) $S_R^{0.5}$ - HV-VV.



(i) $S_R^{0.9}$ - HV-VV.



(j) S_{RV} - HV-VV.

Fonte: Elaborada pelo autor (2023)

4 DISTRIBUIÇÃO BIVARIADA $\mathcal{G}_I^0\text{MB}$ APLICADA ÀS CARACTERÍSTICAS SAR COM BASE EM SPAN

Neste capítulo, está sendo proposta uma nova distribuição bivariada para o atributo (C_{11}, SPAN) . Algumas propriedades matemáticas que tipificam e caracterizam o novo par aleatório são derivadas: vetor de médias, matriz de covariância, função geradora de momentos, esperança condicionais, variância condicionais e momentos condicionais. Além disso, derivam-se as divergências e distâncias de Kullback-Leibler e Rényi para a nova distribuição bivariada, e, em consequência, desenvolvem-se novos testes de hipóteses e novos detectores de mudança para essas medidas de divergência.

Seja a distribuição MBG com parâmetro de locação $\alpha_1, \alpha_2 > 0$ e parâmetro de escala $\gamma > 0$ com fdp definida em (3.7) e seja $Y \sim \Gamma^{-1}(\lambda, \beta)$, $\lambda, \beta > 0$ com densidade

$$f(y; \lambda, \beta) = \frac{\beta^\lambda y^{-\lambda-1}}{\Gamma(\lambda)} \exp\left(-\frac{\beta}{y}\right), \quad (4.1)$$

para $y > 0$.

No contexto de dados SAR, a MM descreve o retorno SAR Z como o produto de duas variáveis aleatórias independentes ($Z = XY$), em que X modela o retroespalhamento (backscatter) e Y modela o ruído speckle. Sendo X uma v.a. com suporte definido nos reais positivos e Y uma variável aleatória escalar (caso monopolarizado), vetorial (caso single-look) ou matricial (multi-look) (HORTA, 2009). Para utilizar a metodologia da MM (ver Capítulo 3) para propor uma nova distribuição bivariada para o par aleatório induzido pelo SPAN (tal como discutido no Capítulo anterior) flexível para outros tipos de terrenos, formula-se a seguinte proposição baseada na distribuição MBG com a distribuição Gama Inversa:

Proposição 4.0.1 *Seja $Y \sim \Gamma^{-1}(\tau, \beta)$ e $\mathbf{x} \sim \text{MBG}(\alpha_1, \alpha_2, \gamma)$ e $\mathbf{x}$ e Y são v.a.s independentes. Então, seja $\mathbf{z} := \mathbf{x} \times Y$ a função densidade conjunta*

$$f_{\mathbf{z}}(z_1, z_2) = \frac{\beta^\tau \Gamma(\tau + \alpha_1 + \alpha_2 + 1) z_1^{\alpha_1-1} (z_2 - z_1)^{\alpha_2-1}}{\gamma^{\alpha_1+\alpha_2} \Gamma(\alpha_1) \Gamma(\alpha_2) \Gamma(\tau) \left(\frac{z_2}{\gamma} + \beta\right)^{\tau+\alpha_1+\alpha_2+1}}, \quad (4.2)$$

para $0 < z_1 < z_2$ e $\beta, \tau > 0$.

A prova está descrita no Apêndice C.

Neste capítulo, iremos apresentar alguns resultados desenvolvidos para o par aleatório com densidade (4.2). Este caso é denotado como $\mathbf{z} \sim \mathcal{G}_I^0\text{MB}(\alpha_1, \alpha_2, \gamma, \beta, \tau)$, em que $(\alpha_1, \alpha_2, \gamma)$ é a parte paramétrica referente ao ruído, enquanto (β, τ) representa a parte referente ao relevo.

Após algumas manipulações algébricas das propriedades matemáticas do $\mathcal{G}_I^0 MB$ (MATHAI; MOSCHOPOULOS, 1991), obtemos o vetor de médias e a matriz de covariância descritos, respectivamente,

$$\mathbb{E}[z] = \frac{\beta \gamma}{\tau - 1} \begin{bmatrix} \alpha_1 \\ \alpha_1 + \alpha_2 \end{bmatrix} \text{ and } \text{Cov}[z] = \frac{\beta^2}{\tau - 1} \left\{ \frac{1}{\tau - 2} \mathbb{E}(x x^\top) - \frac{1}{\tau - 1} \mathbb{E}(x) \mathbb{E}(x^\top) \right\}, \quad (4.3)$$

Em que

$$\mathbb{E}(x x^\top) = \gamma^2 \alpha_1 \begin{bmatrix} 1 + \alpha_1 & 1 + \alpha_1 + \alpha_2 \\ 1 + \alpha_1 + \alpha_2 & (1 + \alpha_2/\alpha_1)(1 + \alpha_1 + \alpha_2) \end{bmatrix}.$$

Da proposição 4.0.1, derivamos as funções marginais z_1 e z_2 e, após algumas manipulações algébricas, encontramos também o coeficiente de correlação entre as variáveis z_1 e z_2 . O seguinte Corolário resume essas propostas teóricas.

Corolário 4.0.2 *Seja $z \sim \mathcal{G}_I^0 MB(\alpha_1, \alpha_2, \gamma, \beta, \tau)$, então as densidades marginais $f_{Z_1}(z_1)$ e $f_{Z_2}(z_2)$ e o coeficiente de correlação são definidos:*

$$f_{Z_1}(z_1) = \frac{\beta^\tau \Gamma(\tau + 1) B(\alpha_2, \tau + \alpha_1 + 1) z_1^{\alpha_1 - 1}}{(\beta \gamma)^{(\tau + \alpha_1 + 1)} \Gamma(\tau) \Gamma(\alpha_1) \Gamma(\alpha_2)} \cdot {}_2F_1\left(\frac{-\alpha_1 - \alpha_2 - \tau}{2}, \frac{-\alpha_1 - \alpha_2 - \tau + 1}{2}; -\alpha_1 - \tau; -\frac{4z_1}{\beta \gamma}\right), \quad (4.4)$$

em que $B(., .)$ é a função beta e ${}_2F_1(., .)$ é a função hipergeométrica,

$$f_{Z_2}(z_2) = \frac{\beta^\tau \Gamma(\tau + \alpha_1 + \alpha_2 + 1) B(\alpha_1, \alpha_2)}{\gamma^{\alpha_1 + \alpha_2} \Gamma(\tau) \Gamma(\alpha_1) \Gamma(\alpha_2)} \frac{z_2^{\alpha_1 + \alpha_2 + 1}}{\left(\frac{z_2}{\gamma} + \beta\right)^{\tau + \alpha_1 + \alpha_2 + 1}} \quad (4.5)$$

$$\text{Cor}(Z_1, Z_2) = \sqrt{\frac{\alpha_1(\alpha_1 + \alpha_2)}{(\tau - \alpha_1 - 1)(\alpha_1 + \alpha_2 + \tau - 1)}}, \quad (4.6)$$

para $\tau > \alpha_1 + 1$, $\text{Cor}(\cdot, \cdot)$ é o operador de correlação.

A prova está descrita no Apêndice D.

Podemos descrever o efeito de τ (parâmetro relacionado com o relevo) e (α_1, α_2) (parâmetro relacionado com o ruído speckle) na estrutura de correlação. Note-se que quanto maior o valor de τ , menor é o coeficiente de correlação. Isso sugere que os valores baixos estão alinhados com cenários de textura mais pronunciada (como áreas florestais urbanas e densas), e os altos indicam áreas mais homogêneas (como pastagens e mares). Além disso, se α_1 ou $\alpha_1 + \alpha_2$ tendem a zero, então a correlação tende a zero. Nesse caso, $\mathbf{x} = (X_1, X_2) \sim MGB(\alpha_1, \alpha_2, \gamma)$ tem ambas as variâncias marginais nulas e o efeito do ruído speckle bidimensional é mínimo. A função geradora do momento (FGM) é definida pelo seguinte Teorema 4.7

Teorema 4.0.3 *Seja $z \sim \mathcal{G}_I^0 MB(\boldsymbol{\theta})$. Então, a função geradora de momentos de z é definida como*

$$M_z(t) = \frac{\tau}{\beta} {}_3F_4\left(\alpha_1, \frac{\alpha_1 + \alpha_2}{2}, \frac{\alpha_1 + \alpha_2 + 1}{2}; 1, \alpha_1 + \alpha_2, \frac{1 - \tau}{2}, \frac{-\tau}{2}; t_1 t_2 \gamma^2 \beta^2\right) \quad (4.7)$$

A prova está descrita no Apêndice E.

No estudo aplicado deste capítulo, assume-se que a distribuição para modelar o relevo tem média unitária (ou seja, $\beta := \tau + 1$) e a densidade (4.2) colapsa em

$$f_{z^*}(z_1, z_2) = \frac{\beta^\beta \Gamma(\beta + \alpha_1 + \alpha_2)}{\gamma^{\alpha_1 + \alpha_2} \Gamma(\beta) \Gamma(\alpha_1) \Gamma(\alpha_2)} \frac{z_1^{\alpha_1 - 1} (z_2 - z_1)^{\alpha_2 - 1}}{\left(\frac{z_2}{\gamma} + \beta\right)^{\beta + \alpha_1 + \alpha_2}}. \quad (4.8)$$

4.1 ESTIMATIVA DE MÁXIMA VEROSSIMILHANÇA PARA OS PARÂMETROS $\mathcal{G}_I^0 MB$

Agora considere derivar estimativas de máxima verossimilhança (EMVs) para $\boldsymbol{\theta} = (\gamma, \alpha_1, \alpha_2, \beta, \tau)$. Sejam $(z_{11}, z_{21}), \dots, (z_{1n}, z_{2n})$ uma amostra observada obtida de uma amostra aleatória, $\{\dots, (Z_{1i}, Z_{2i}), \dots\}$, de $\mathbf{z} = (Z_1, Z_2) \sim \mathcal{G}_I^0 MB(\boldsymbol{\theta})$. Assim, a função log-verossimilhança associada em $\boldsymbol{\theta}$ é

$$\begin{aligned} \ell(\boldsymbol{\theta}) := & \sum_{i=1}^n \log f_{\mathbf{z}}(z_{1i}, z_{2i}) = n \log \Gamma(\tau + \alpha_1 + \alpha_2 + 1) + n\tau \log \beta \\ & - n \log \Gamma(\tau) - n \log \Gamma(\alpha_1) - n \log \Gamma(\alpha_2) \\ & - n(\alpha_1 + \alpha_2) \log \gamma + (\alpha_1 - 1) \sum_{i=1}^n \log z_{1i} + (\alpha_2 - 1) \sum_{i=1}^n \log(z_{2i} - z_{1i}) \\ & - (\tau + \alpha_1 + \alpha_2 + 1) \sum_{i=1}^n \log \left(\frac{z_{2i}}{\gamma} + \beta \right). \end{aligned}$$

Os EMVs para $\boldsymbol{\theta}$, seja $\hat{\boldsymbol{\theta}}$, são definidos como $\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta} \in \Theta} [\ell(\boldsymbol{\theta})]$ tal que Θ é o espaço paramétrico associado ou, equivalentemente, EMVs representam uma solução no seguinte sistema não linear

$$\left[\frac{\partial \ell(\boldsymbol{\theta})}{\partial \gamma}, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha_1}, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha_2}, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \beta}, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \tau} \right] \bigg|_{\gamma=\hat{\gamma} \alpha_1=\hat{\alpha}_1 \alpha_2=\hat{\alpha}_2 \beta=\hat{\beta} \tau=\hat{\tau}} = [0, 0, 0, 0, 0]$$

$$\iff \begin{cases} \frac{\sum_{i=1}^n \log(z_{2i} + \hat{\gamma}\hat{\beta}) - \sum_{i=1}^n \log x_{1i}}{n} = \Psi(\hat{\tau} + \hat{\alpha}_1 + \hat{\alpha}_2 + 1) - \Psi(\hat{\alpha}_1), \\ \frac{\sum_{i=1}^n \log(z_{2i} + \hat{\gamma}\hat{\beta}) - \sum_{i=1}^n \log(x_{2i} - x_{1i})}{n} = \Psi(\hat{\tau} + \hat{\alpha}_1 + \hat{\alpha}_2 + 1) - \Psi(\hat{\alpha}_2), \\ \frac{\sum_{i=1}^n \log(z_{2i} + \hat{\gamma}\hat{\beta})}{n} = \log \hat{\gamma} + \log \hat{\beta} - \Psi(\hat{\tau}) + \Psi(\hat{\tau} + \hat{\alpha}_1 + \hat{\alpha}_2 + 1), \\ \sum_{i=1}^n \log(z_{2i} + \hat{\gamma}\hat{\beta}) = \frac{\hat{\gamma}\hat{\beta}(\hat{\tau} + \hat{\alpha}_1 + \hat{\alpha}_2 + 1)}{n\hat{\tau}}, \\ \sum_{i=1}^n \log(z_{2i} + \hat{\gamma}\hat{\beta}) = \frac{\gamma(\hat{\tau} + \hat{\alpha}_1 + \hat{\alpha}_2 + 1)}{n(\hat{\alpha}_1 + \hat{\alpha}_2)}, \end{cases}$$

em que $\Psi(\cdot)$ é a função Psi.

Na seção de resultados numéricos, será verificado o desempenho deste procedimento. A partir de agora, também usaremos o estimador de máxima verossimilhança (EMV) para $\hat{\boldsymbol{\theta}}$ para fins de teoria assintótica.

4.2 ESPERANÇA E MOMENTOS CONDICIONAIS

Nesta seção derivam-se os resultados analíticos para os valores esperados condicionais do tipo $\mathbb{E}(Z_1 | Z_2)$ e $\mathbb{E}(Z_2 | Z_1)$, as variâncias condicionais do tipo $\text{Var}(Z_1 | Z_2)$ e $\text{Var}(Z_2 | Z_1)$ e os momentos condicionais do tipo $M_{Z_1|Z_2}(t)$ e $M_{Z_2|Z_1}(t)$. Portanto, apresentaremos estes resultados em dois corolários, no primeiro analisamos a influência de z_1 sobre z_2 e no segundo a influência de z_2 sobre z_1 .

Corolário 4.2.1 *Seja $z \sim \mathcal{G}_I^0 MB(\theta)$, então a esperança condicional do tipo $\mathbb{E}(Z_1 | Z_2)$, a variância condicional do tipo $\text{Var}(Z_1 | Z_2)$ e o momento condicional do tipo $M_{Z_1|Z_2}(t)$ são definidos como:*

$$\mathbb{E}(Z_1 | Z_2) = \frac{B(\alpha_1 + 1, \alpha_2) z_2}{B(\alpha_1, \alpha_2)}, \quad (4.9)$$

$$\text{Var}(Z_1 | Z_2) = \left[\frac{B(\alpha_1, \alpha_2) B(\alpha_1 + 2, \alpha_2) - B(\alpha_1 + 1, \alpha_2)^2}{B(\alpha_1, \alpha_2)^2} \right] z_2^2 \quad (4.10)$$

$$M_{Z_1|Z_2}(t) = {}_2F_2\left(\frac{\alpha_1 + 1}{2}, \frac{\alpha_1}{2}; \frac{\alpha_1 + \alpha_2}{2}, \frac{\alpha_1 + \alpha_2 + 1}{2}; z_2\right) \quad (4.11)$$

A prova desse Corolário 4.2.1 está descrita no Apêndice F.

Corolário 4.2.2 *Seja $z \sim \mathcal{G}_I^0 MB(\theta)$, então a esperança condicional do tipo $\mathbb{E}(Z_2 | Z_1)$, a variância condicional do tipo $\text{Var}(Z_2 | Z_1)$ e o momento condicional do tipo $M_{Z_2|Z_1}(t)$ são definidos como:*

$$\mathbb{E}(Z_2 | Z_1) = \frac{(-\beta\gamma) \alpha_2}{\tau + \alpha_1} \frac{{}_2F_1\left(-\alpha_2, \tau + \alpha_1; -\alpha_2 - 1; \frac{z_1}{\beta\gamma}\right)}{{}_2F_1\left(\delta_1, \delta_2; \delta_3; \frac{-4z_1}{\beta\gamma}\right)}, \quad (4.12)$$

$$\begin{aligned}
 \text{Var}(Z_2 | Z_1) = & \frac{(\beta\gamma\alpha_2)^2 {}_2F_1\left(-\alpha_2 - 1, \tau + \alpha_1 - 1; -\alpha_2 - 2; \frac{z_1}{\beta\gamma}\right)}{(\tau + \alpha_1)(\tau + \alpha_1 - 1) {}_2F_1\left(\delta_1, \delta_2; \delta_3; \frac{-4z_1}{\beta\gamma}\right)} \\
 & - \left[\frac{(-\beta\gamma)\alpha_2 {}_2F_1\left(-\alpha_2, \tau + \alpha_1; -\alpha_2 - 1; \frac{z_1}{\beta\gamma}\right)}{\tau + \alpha_1} \frac{{}_2F_1\left(\delta_1, \delta_2; \delta_3; \frac{-4z_1}{\beta\gamma}\right)}{{}_2F_1\left(\delta_1, \delta_2; \delta_3; \frac{-4z_1}{\beta\gamma}\right)} \right]^2
 \end{aligned} \tag{4.13}$$

$$M_{Z_1|Z_2}(t) = -\frac{\tau + \alpha_1 + 1}{(\alpha_2 - 1)(-\beta\gamma)^{\alpha_2+1}} \frac{{}_1F_1(-\alpha_1 - 1; 1; -tz_1)}{{}_2F_1\left(\delta_1, \delta_2; \delta_3; \frac{-4z_1}{\beta\gamma}\right)} \tag{4.14}$$

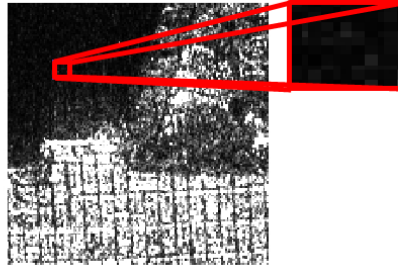
A prova desse Corolário 4.2.2 está descrita no Apêndice G.

4.3 APLICAÇÃO A DADOS POLSAR COM BASE NO SPAN PARA $\mathcal{G}_I^0\text{MB}$.

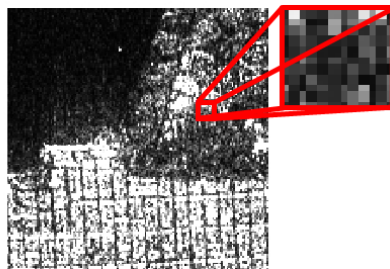
Nesta seção, descreve-se o comportamento da $G_I^0(\alpha_1, \alpha_2, \gamma, \beta, \tau)$ através do seu caso particular descrito em 4.8. Portanto, discutiremos o efeito do parâmetro β sobre a fdp, ou seja, descrever a influência do relevo sobre uma imagem PolSAR através de ilustrações de curvas de níveis por EMVs. Apresentaremos esse estudo por meio dos modelos de curvas teóricas e empíricas. Os dados empíricos se referem à cidade de São Francisco, Califórnia-USA, os quais foram capturados por AIRSAR. Um exemplo que descreve a dispersão de valores entre o canal de intensidade HH e o par de valor (HH, HH+VV) pode ser visto na figura 11. Para fins ilustrativos, foram extraídas amostras de tamanho $N = 25$ para descrever as regiões oceânicas (Cena1), florestais (Cena2) e urbanas (Cena 3).

As figuras 13d - 13i ilustram o comportamento das curvas teóricas(vermelha) e empíricas (preta) sobre o conjunto de dados PolSAR (azul) para valores de $\beta = 2.5$. Foi utilizado esse valor para verificar a influência do relevo nos diferentes tipos de terrenos. De modo geral, as curvas estimadas obtiveram ótimos ajustes aos dados, com destaque para a região oceânica os pares HH com HH-VV (ver figura 13b), em regiões florestais foram os pares HH com HH-HV (ver figura 13d) e, por fim, na região urbana, foram os pares HV com HV-VV (ver figura 13i).

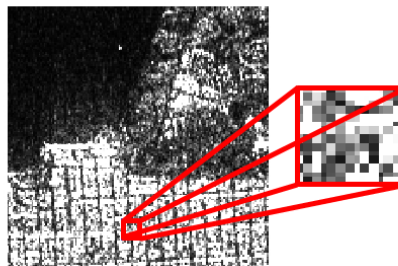
Figura 11 – Cenários de simulação obtida da região de São Francisco (USA) e gráfico de dispersão de valores de pares (HH, HH+VV).



(a) Cenário 1.



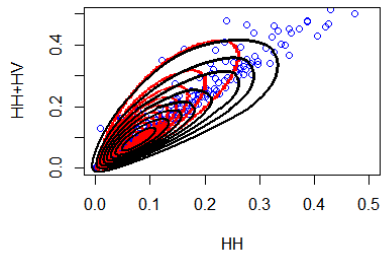
(b) Cenário 2.



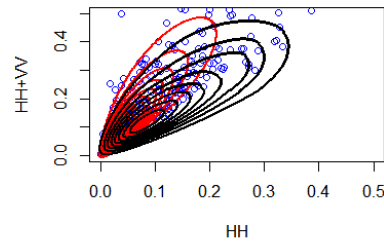
(c) Cenário 3.

Fonte: Elaborada pelo autor (2023)

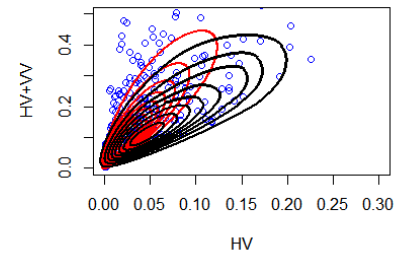
Figura 12 – Curvas de níveis teóricas e empíricas ajustadas a fdp $\mathcal{G}_I^0$ MB.



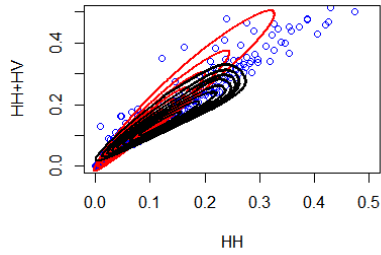
(a) Cenário 1.



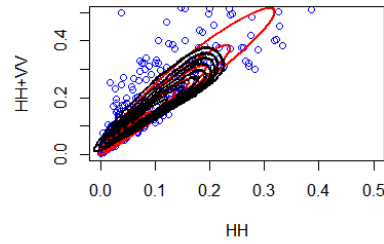
(b) Cenário 1.



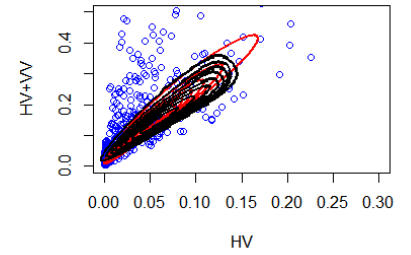
(c) Cenário 1.



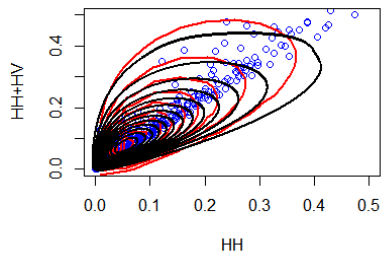
(d) Cenário 2.



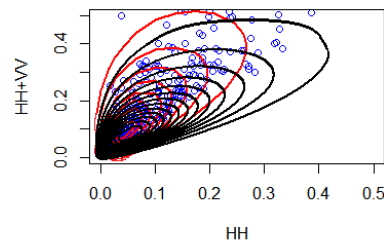
(e) Cenário 2.



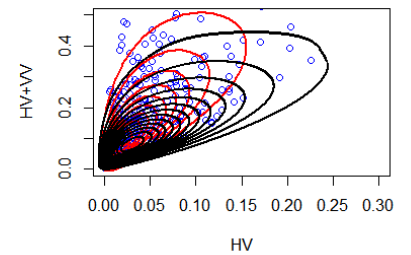
(f) Cenário 2.



(g) Cenário 3.



(h) Cenário 3.



(i) Cenário 3.

Fonte: Elaborada pelo autor (2023)

4.4 TESTE DE HIPÓTESE BASEADOS NA ESTATÍSTICA DA RAZÃO ENTRE VEROSIMILHANÇAS PARA A $\mathcal{G}_I^0 MB$

Sejam $\boldsymbol{\theta}_1 = (\alpha_{11}, \alpha_{12}, \gamma_1, \beta_1, \tau_1)$ e $\boldsymbol{\theta}_2 = (\alpha_{21}, \alpha_{22}, \gamma_2, \beta_2, \tau_2)$ vetores de parâmetros de dois pares aleatórios, digamos de duas amostras aleatórias de $\mathcal{G}_I^0 MB(\alpha_1 1, \alpha_2 1, \gamma_1, \beta_1, \tau_1)$ e $\mathcal{G}_I^0 MB(\alpha_1 2, \alpha_2 2, \gamma_2, \beta_2, \tau_2)$. Admita ainda que $[(X_{11}^{(1)}, X_{21}^{(1)}), \dots, (X_{1M_1}^{(1)}, X_{2M_1}^{(1)})]$ e

$[(X_{12}^{(1)}, X_{22}^{(1)}), \dots, (X_{1M_2}^{(1)}, X_{2M_2}^{(1)})]$ são duas amostras aleatórias obtidas de $\boldsymbol{x}$ e $\boldsymbol{y}$. A estatística RV é definida como

$$S_{RV} = -2 \log \frac{\sup_{\boldsymbol{\theta} \in \Theta_0} \ell(\boldsymbol{\theta})}{\sup_{\boldsymbol{\theta} \in \Theta} \ell(\boldsymbol{\theta})}, \quad (4.15)$$

em que $\Theta = \Theta_0 \cup \Theta_1$ e $\Theta_0 \cap \Theta_1 = \emptyset$.

Assim, temos,

$$\begin{aligned} \ell(\boldsymbol{\theta} \in \Theta_0) &= (M_1 + M_2) \log(\vartheta) + (\alpha_1 - 1) \left[\sum_{i=1}^{M_1} \log X_{1i}^{(1)} + \sum_{j=1}^{M_2} \log X_{1j}^{(2)} \right] \\ &+ (\alpha_2 - 1) \left[\sum_{i=1}^{M_1} \log(X_{2i}^{(1)} - X_{1i}^{(1)}) + \sum_{j=1}^{M_2} \log(X_{2j}^{(2)} - X_{1j}^{(2)}) \right] \\ &- (\tau + \alpha_1 + \alpha_2 + 1) \left[\sum_{i=1}^{M_1} \log\left(\frac{X_{2i}^{(1)}}{\gamma} + \beta\right) + \sum_{j=1}^{M_2} \log\left(\frac{X_{2j}^{(2)}}{\gamma} + \beta\right) \right] \end{aligned}$$

em que $\vartheta = \frac{\Gamma(\tau + \alpha_1 + \alpha_2 + 1)\beta^\tau}{\Gamma(\tau)\Gamma(\alpha_1)\Gamma(\alpha_2)\gamma^{\alpha_1 + \alpha_2}}$

$$\begin{aligned} \ell(\boldsymbol{\theta} \in \Theta) &= M_1 \log(\vartheta_1) + M_2 \log(\vartheta_2) + (\alpha_{11} - 1) \sum_{i=1}^{M_1} \log X_{1i}^{(1)} + (\alpha_{12} - 1) \sum_{j=1}^{M_2} \log X_{1j}^{(2)} \\ &+ (\alpha_{21} - 1) \sum_{i=1}^{M_1} \log(X_{2i}^{(1)} - X_{1i}^{(1)}) + (\alpha_{22} - 1) \sum_{j=1}^{M_2} \log(X_{2j}^{(2)} - X_{1j}^{(2)}) \\ &- (\tau_1 + \alpha_1 1 + \alpha_2 1 + 1) \sum_{i=1}^{M_1} \log\left(\frac{X_{2i}^{(1)}}{\gamma_1} + \beta_1\right) \\ &- (\tau_2 + \alpha_1 2 + \alpha_2 2 + 1) \sum_{j=1}^{M_2} \log\left(\frac{X_{2j}^{(2)}}{\gamma_2} + \beta_2\right), \end{aligned}$$

Em que $\vartheta_i = \frac{\Gamma(\tau_i + \alpha_{1i} + \alpha_{2i} + 1)\beta_i^{\tau_i}}{\Gamma(\tau_i)\Gamma(\alpha_{1i})\Gamma(\alpha_{2i})\gamma_i^{\alpha_{1i} + \alpha_{2i}}}$ para $i = 1, 2$, α_{1i} para $i = 1, 2$ são os parâmetros de localização para a primeira distribuição considerada, α_{2i} para $i = 1, 2$ são os parâmetros de localização para a segunda distribuição considerada, β_i para $i = 1, 2$ são os parâmetros de localização para i-ésima distribuição, γ_i para $i = 1, 2$ é o parâmetro de escala para distribuição i-ésima e τ_i para $i = 1, 2$ é o parâmetro de escala para distribuição i-ésima.

4.5 NOVOS TESTES DE HIPÓTESES BASEADOS EM DIVERGÊNCIAS PARA A DISTRIBUIÇÃO BIVARIADA $\mathcal{G}_I^0\text{MB}$

Os testes de hipóteses baseados em divergências são uma ferramenta importante na análise estatística e podem ser usados para testar uma ampla variedade de hipóteses sobre distribuições de probabilidade. A escolha da medida de divergência adequada depende das suposições específicas do problema em questão e deve ser feita com cuidado, para garantir a validade e a robustez dos resultados do teste.

Em termos de aplicação na classificação e segmentação de imagens PolSAR, a divergência KL pode ser usada para medir a diferença entre as distribuições de intensidade de diferentes tipos de objetos em uma imagem PolSAR, por exemplo, para medir a diferença entre as distribuições de intensidade de alvos urbanos e não urbanos em uma imagem PolSAR (FANG; ZHU, 2020), ou para medir a diferença entre as distribuições de intensidade de diferentes tipos de vegetação (WANG et al., 2019).

Teorema 4.5.1 *Sejam $\theta_1 = (\alpha_{11}, \alpha_{12}, \gamma_1, \beta_1, \tau_1)$ e $\theta_2 = (\alpha_{21}, \alpha_{22}, \gamma_2, \beta_2, \tau_2)$ vetores de parâmetros de dois pares aleatórios, digamos x e y , seguindo a distribuição de $\mathcal{G}_I^0\text{MB}$. A divergência Kullback-Leibler entre x e y é*

$$\begin{aligned}
 D_{KL}(x||y) = D_{KL}(\theta_1||\theta_2) = & \frac{\tau_1}{\beta_1} \left\{ (\alpha_{11} - \alpha_{21}) [\Psi(\alpha_{11}) + \log \beta_1 \gamma_1 - \Psi(\tau_1 + 1)] + \right. \\
 & (\alpha_{12} - \alpha_{22}) [\Psi(\alpha_{12}) + \log \beta_1 \gamma_1 - \Psi(\tau_1 + 1)] + \\
 & \log \left(\frac{\beta_1^{\tau_1} \Gamma(\tau_1 + \alpha_{11} + \alpha_{12} + 1)}{\Gamma(\tau_1) \Gamma(\alpha_{11}) \Gamma(\alpha_{12}) \gamma_1^{\alpha_{11} + \alpha_{12}}} \cdot \frac{\Gamma(\tau_2) \Gamma(\alpha_{21}) \Gamma(\alpha_{22}) \gamma_2^{\alpha_{21} + \alpha_{22}}}{\beta_2^{\tau_2} \Gamma(\tau_2 + \alpha_{21} + \alpha_{22} + 1)} \right) + \\
 & (\tau_2 + \alpha_{21} + \alpha_{22} + 1) \left[\Psi(\tau_1 + \alpha_{11} + \alpha_{12} + 1) + \log \beta_2 - \Psi(\tau_1 + 1) - \right. \\
 & \frac{(\alpha_{11} + \alpha_{12}) \left(1 - \frac{\beta_1 \gamma_1}{\beta_2 \gamma_2} \right)}{\tau_1 + \alpha_{11} + \alpha_{12} + 1} {}_3F_2 \left(\alpha_{11} + \alpha_{12} + 1, 1, 1; 2, \tau_1 + \alpha_{11} + \alpha_{12} + 2; 1 - \frac{\beta_1 \gamma_1}{\beta_2 \gamma_2} \right) \left. \right] - \\
 & \left. (\tau_1 + \alpha_{11} + \alpha_{12} + 1) [\Psi(\tau_1 - 1) + \Psi(\tau_1 + \alpha_{11} + \alpha_{12} + 1) + \log \beta_1] \right\}, \quad (4.16)
 \end{aligned}$$

em que ${}_3F_2(\cdot, \cdot, \cdot; \cdot, \cdot; \cdot)$ é a função hipergeométrica generalizada, α_{1i} para $i = 1, 2$ são os parâmetros de localização para a primeira distribuição considerada, α_{2i} para $i = 1, 2$ são os parâmetros de localização para a segunda distribuição considerada, β_i para $i = 1, 2$, são os parâmetros de localização para i -ésima distribuição, γ_i para $i = 1, 2$, é o parâmetro de escala

para distribuição i -ésima e τ_i para $i = 1, 2$, é o parâmetro de escala para distribuição i -ésima.

A prova está descrita no Apêndice H. No que se segue, apresentamos a distância KL associada.

Corolário 4.5.2 *Sejam $\theta_1 = (\alpha_{11}, \alpha_{12}, \gamma_1, \beta_1, \tau_1)$ e $\theta_2 = (\alpha_{21}, \alpha_{22}, \gamma_2, \beta_2, \tau_2)$, no contexto do Teorema 4.5.1, temos*

$$\begin{aligned}
 d_{KL}(\theta_1, \theta_2) = & \frac{1}{2} \left\{ \left(\frac{\tau_2}{\beta_2} - \frac{\tau_1}{\beta_1} \right) \left\{ \log \Gamma(\eta_1) - \log \Gamma(\tau_1) - \log \Gamma(\alpha_{11}) - \log \Gamma(\alpha_{12}) + \right. \right. \\
 & \log \Gamma(\tau_2) + \log \Gamma(\alpha_{21}) + \log \Gamma(\alpha_{22}) - \log \Gamma(\eta_2) + \log \beta_2 - \log \beta_1 + \\
 & (\alpha_{11} - \alpha_{21}) [\Psi(\alpha_{11}) - \log \beta_1 \gamma_1 - \log \beta_2 \gamma_2 - \Psi(\eta_1) - \Psi(\alpha_{21}) + \Psi(\eta_2)] + \\
 & (\alpha_{12} - \alpha_{22}) [\Psi(\alpha_{12}) - \log \beta_1 \gamma_1 - \log \beta_2 \gamma_2 - \Psi(\eta_1) - \Psi(\alpha_{22}) + \Psi(\eta_2)] + \\
 & \left. (\tau_2 - \tau_1) [\Psi(\eta_1) - \Psi(\tau_1 + 1) - \Psi(\eta_2) + \Psi(\tau_2 + 1)] \right\} - \\
 & \frac{\tau_1}{\beta_1} \left[(\alpha_{11} + \alpha_{12}) \left(1 - \frac{\beta_1 \gamma_1}{\beta_2 \gamma_2} \right) {}_3F_2 \left(\alpha_{11} + \alpha_{12} + 1, 1, 1; 2, \eta_1; 1 - \frac{\beta_1 \gamma_1}{\beta_2 \gamma_2} \right) \right] - \\
 & \frac{\tau_2}{\beta_2} \left[(\alpha_{21} + \alpha_{22}) \left(1 - \frac{\beta_2 \gamma_2}{\beta_1 \gamma_1} \right) {}_3F_2 \left(\alpha_{21} + \alpha_{22} + 1, 1, 1; 2, \eta_2; 1 - \frac{\beta_2 \gamma_2}{\beta_1 \gamma_1} \right) \right] \Big\}, \quad (4.17)
 \end{aligned}$$

em que $\eta_i = \tau_i + \alpha_{i1} + \alpha_{i2} + 1$ para $i = 1, 2$.

No teorema 4.5.3 encontra-se a expressão de divergência de Rényi para a distribuição $\mathcal{G}_I^0 \text{MB}$.

Teorema 4.5.3 *Sejam $\theta_1 = (\alpha_{11}, \alpha_{12}, \gamma_1, \beta_1, \tau_1)$ e $\theta_2 = (\alpha_{21}, \alpha_{22}, \gamma_2, \beta_2, \tau_2)$ vetores de parâmetros de dois pares aleatórios, digamos x e y , seguindo a distribuição de $\mathcal{G}_I^0 \text{MB}$. A divergência de Rényi entre x e y é*

$$\begin{aligned}
 D_{R:r}(\theta_1 || \theta_2) = & \frac{1}{r(1-r)} \left\{ r \log A_1 + (1-r) \log A_2 - r \kappa_1 \log \beta_1 - (1-r) \kappa_2 \log \beta_2 \right. \\
 & + (\eta_1 + \eta_2) \log(\gamma_1 \beta_1) + \log \text{Beta}(\eta_1, \eta_2) + \log \text{Beta}(\eta_1 + \eta_2, r\tau_1 + (1-r)\tau_2 + 1) \\
 & \left. + \log {}_2F_1 \left(\eta_1 + \eta_2, (1-r)\kappa_2; r\kappa_1 + (1-r)\kappa_2; 1 - \frac{\beta_1 \gamma_1}{\beta_2 \gamma_2} \right) \right\}, \quad (4.18)
 \end{aligned}$$

em que $A_i := [\beta_i^{\tau_i} \Gamma(\tau_i + \alpha_{i1} + \alpha_{i2} + 1)] [\Gamma(\tau_i) \Gamma(\alpha_{i1}) \Gamma(\alpha_{i2}) \gamma_i^{\alpha_{i1} + \alpha_{i2}}]^{-1}$ para $i = 1, 2$, $\eta_1 = \alpha_{21} + r \cdot \alpha_{11} - r \cdot \alpha_{21}$, $\eta_2 = \alpha_{22} + r \cdot \alpha_{12} - r \cdot \alpha_{22}$, $\kappa_1 = \tau_1 + \alpha_{11} + \alpha_{12} + 1$ e $\kappa_2 = \tau_2 + \alpha_{21} + \alpha_{22} + 1$.

A prova está descrita no Apêndice I. No que se segue, apresentamos a distância Rényi associada.

Corolário 4.5.4 *Sejam $\theta_1 = (\alpha_{11}, \alpha_{12}, \gamma_1, \beta_1, \tau_1)$ e $\theta_2 = (\alpha_{21}, \alpha_{22}, \gamma_2, \beta_2, \tau_2)$. Sejam $\eta_1 = \alpha_{11} + r \cdot \alpha_{11} - r \cdot \alpha_{21}$, $\eta_2 = \alpha_{12} + r \cdot \alpha_{12} - r \cdot \alpha_{22}$ no contexto do Teorema 4.5.3, temos*

$$\begin{aligned} d_{R:r}(\theta_1 || \theta_2) = & -\frac{\log 2}{r(r-1)} + \frac{1}{r(r-1)} \log \left\{ A_1^r A_2^{1-r} \beta_1^{-r \kappa_1} \beta_2^{-(1-r) \kappa_2} \text{Beta}(\eta_1, \eta_2) \right. \\ & \times (\beta_1 \gamma_1)^{\eta_1 + \eta_2} \text{Beta}(\eta_1, \eta_2) \text{Beta}(\eta_1 + \eta_2, r\tau_1 + (1-r)\tau_2 + 1) \\ & \times {}_2F_1 \left(\eta_1 + \eta_2, (1-r)\kappa_2; r\kappa_1 + (1-r)\kappa_2; 1 - \frac{\beta_1 \gamma_1}{\beta_2 \gamma_2} \right) + A_2^r A_1^{1-r} \beta_2^{-r \kappa_2} \beta_1^{-(1-r) \kappa_1} \\ & \times (\beta_2 \gamma_2)^{\tilde{\eta}_1 + \tilde{\eta}_2} \text{Beta}(\tilde{\eta}_1, \tilde{\eta}_2) \text{Beta}(\tilde{\eta}_1 + \tilde{\eta}_2, r\tau_2 + (1-r)\tau_1 + 1) \\ & \left. \times {}_2F_1 \left(\tilde{\eta}_1 + \tilde{\eta}_2, (1-r)\kappa_1; r\kappa_2 + (1-r)\kappa_1; 1 - \frac{\beta_2 \gamma_2}{\beta_1 \gamma_1} \right) \right\}, \end{aligned} \quad (4.19)$$

em que $\tilde{\eta}_1 = \alpha_{11} + r \cdot \alpha_{21} - r \cdot \alpha_{11}$, $\tilde{\eta}_2 = \alpha_{12} + r \cdot \alpha_{22} - r \cdot \alpha_{12}$.

Segue-se do Lema 2.1.1 e a Proposição 2.1.2 (tomando $M = 5$), combinado com os Teoremas 4.5.1 e 4.5.3 e os Corolários 4.5.2 e 4.5.4, dois novos testes de hipóteses, para a distribuição $\mathcal{G}_I^0 \text{MB}$.

- Estatística de Kullback–Leibler - $S_{KL}(\cdot, \cdot)$:

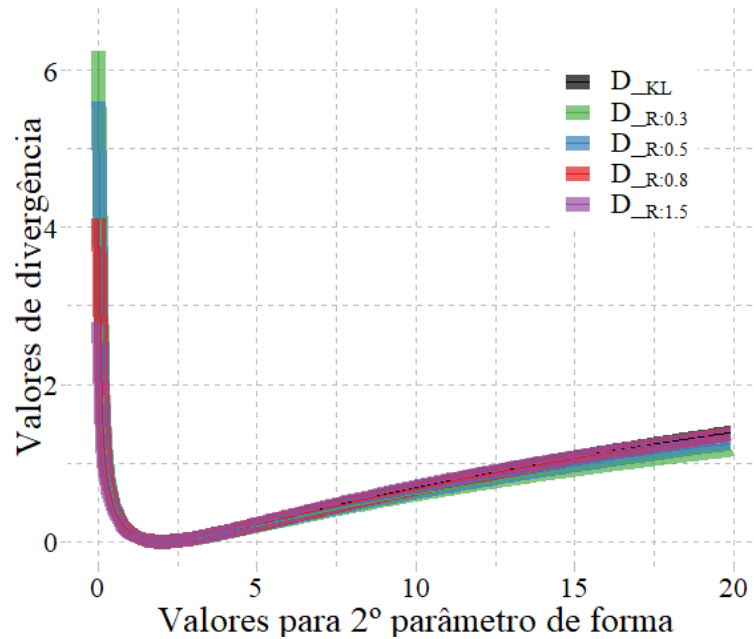
$$\begin{aligned} S_{KL}([\hat{\alpha}_{11}, \hat{\alpha}_{12}, \hat{\gamma}_1, \hat{\beta}_1, \hat{\lambda}_1], [\hat{\alpha}_{21}, \hat{\alpha}_{22}, \hat{\gamma}_2, \hat{\beta}_2, \hat{\lambda}_2]) \\ = \frac{2mn}{m+n} d_{KL}([\hat{\alpha}_{11}, \hat{\alpha}_{12}, \hat{\gamma}_1, \hat{\beta}_1, \hat{\lambda}_1], [\hat{\alpha}_{21}, \hat{\alpha}_{22}, \hat{\gamma}_2, \hat{\beta}_2, \hat{\lambda}_2]) \\ \xrightarrow[m, n \rightarrow \infty]{\mathcal{D}} \chi_5^2. \end{aligned}$$

- Estatística de Rényi de ordem r $S_R^r(\cdot, \cdot)$:

$$\begin{aligned} S_R^r([\hat{\alpha}_{11}, \hat{\alpha}_{12}, \hat{\gamma}_1, \hat{\beta}_1, \hat{\lambda}_1], [\hat{\alpha}_{21}, \hat{\alpha}_{22}, \hat{\gamma}_2, \hat{\beta}_2, \hat{\lambda}_2]) \\ = \frac{2mn}{m+n} d_R^r([\hat{\alpha}_{11}, \hat{\alpha}_{12}, \hat{\gamma}_1, \hat{\beta}_1, \hat{\lambda}_1], [\hat{\alpha}_{21}, \hat{\alpha}_{22}, \hat{\gamma}_2, \hat{\beta}_2, \hat{\lambda}_2]) \\ \xrightarrow[m, n \rightarrow \infty]{\mathcal{D}} \chi_5^2. \end{aligned}$$

A figura 13 mostra curvas da divergência derivada $D_{\bullet}(\theta_1 || \theta_2(x))$ para a distribuição (4.2), em que $\theta_1(x) = (1, 1, 2, 1)$ e $\theta_2(x) = (1, 1, x, 1)$ e x varia em torno de $(0, 20)$. É perceptível que a divergência Rényi converge para a divergência KL quando r tende a um, como esperado.

Figura 13 – Medidas de distância entre duas variáveis aleatórias $\mathcal{G}_I^0\text{MB}$ em função de α_1 .



Fonte: Elaborada pelo autor (2023)

Visando compreender melhor os resultados teóricos obtidos nessa seção, a seguir será realizado um estudo de simulação (ver seção 4.6) que ajudará a fornecer uma compreensão mais profunda dos testes e de como ele deve ser aplicado. Teremos outra seção (ver seção 4.7) cuja ênfase será na aplicação prática do algoritmo, mostrando como ele pode ser usado em situações reais, especificamente, para resolver problemas de detecção de mudanças em dados reais.

4.6 ESTUDO DE SIMULAÇÃO

Nesta seção, realiza-se um estudo de simulação para avaliar o comportamento dos testes propostos na seção 4.5. Inicialmente quantifica-se o desempenho das ferramentas estatísticas propostas para construir detectores de mudança para as imagens PolSAR com base em testes de duas amostras, baseados nas divergências Kullback-Leibler, Rényi (de ordem $\beta = 0.2, 0.5, 0.9$ e 1.5) e estatística da razão entre verossimilhanças.

Os métodos KL, Rényi e RV foram executados com 1000 iterações de Monte Carlo e com tamanhos de amostra $N = 9, 25, 49, 81, 121, 225$ e 441 . Como figura de mérito, usamos os tamanhos de teste empíricos nos níveis nominais 1%, 5% e 10%. Nesse estudo, utilizaremos a fdp descrita em 4.8; sobre os parâmetros em $H0 : (\alpha_{11}; \alpha_{12}; \gamma_1; \beta_1) = (\alpha_{21}; \alpha_{22}; \gamma_2; \beta_2)$ serão definidos nas estatísticas de teste usando MLEs. Com relação aos valores iniciais dos parâmetros, neste estudo adotaremos $\beta = (1.5, 2$ e $2.5)$, e os demais valores dos parâmetros são descritos na Tabela 1, página 36.

Nas Tabelas 5, 6 e 7, estão os resultados dos tamanhos dos testes empíricos, ou seja, as taxas de rejeição sob a hipótese nula dos testes propostos em níveis nominais de 1%, 5% e 10%. Os melhores resultados são destacados em cinza. Em geral, pode-se ver que o teste baseado em S_{KL} apresentou o melhor resultado, visto que seus valores foram os mais próximos aos níveis nominais assumidos. O teste baseado na estatística $S_R^{1.5}$ teve um péssimo desempenho. Quando comparado com o S_{RV} , os testes baseados nas divergências KL e Rényi obtiveram ótimos resultados com valores próximos aos níveis nominais adotados.

Na Tabela 5, podemos destacar o desempenho dos seguintes testes propostos por região:

- No nível nominal (10%), apenas o método $S_R^{0.9}$ se destaca para o primeiro cenário; para o segundo cenário S_{KL} e $S_R^{0.5}$ foram obtidos os melhores resultados,
- Em relação ao nível (5%), o método $S_R^{0.5}$ se destacou no primeiro e terceiro cenários, e no segundo cenário somente o método $S_R^{1.5}$. Entretanto, os métodos S_{KL} , $S_R^{0.2}$ e $S_R^{0.9}$ obtiveram péssimas estimativas.
- No nível 1%, os métodos $S_R^{0.2}$, $S_R^{0.5}$ e $S_R^{0.9}$ tiveram um desempenho ótimo para o terceiro cenário quando $N = 121$ e 441 , S_{KL} e $S_R^{0.2}$ superaram no segundo cenário quando $N = 225$.

Na Tabela 6, podemos destacar o desempenho dos seguintes testes propostos por região:

- No nível nominal de 10%, os seguintes métodos S_{KL} , $S_R^{0.2}$, $S_R^{0.5}$ e $S_R^{0.9}$ obtiveram desempenhos ótimos, para o segundo cenário quando $N = 441$ e para o terceiro cenário quando $N = 49$; 225 e 441.
- Com relação ao nível de 5%, os métodos $S_R^{0.2}$, $S_R^{0.5}$ e $S_R^{0.9}$ quando $N = 441$, no terceiro cenário, apenas S_{KL} quando $N = 225$, no primeiro cenário nenhum método gerou boas estimativas.
- Por fim, no nível de 1% todos os métodos tiveram bom desempenho no segundo cenário, quando $N = 441$, no primeiro cenário somente o método S_{KL} e no terceiro $S_R^{0.2}$ método.

Na Tabela 7, podemos destacar o desempenho dos seguintes testes propostos por região:

- Para os níveis 10% e 5%, no primeiro e segundo cenários, o método $S_R^{0.9}$, seguido pelo método $S_R^{1.5}$ quando $N = 441$ e $S_R^{0.2}$ quando $N = 225$, respectivamente.
- No nível nominal de 1%, apenas o terceiro cenário obteve boas estimativas para os métodos S_{KL} , $S_R^{0.2}$ e $S_R^{0.9}$.

Tabela 5 – Taxas de Rejeição de Testes de Divergência sob $H_0 : (\alpha_{11}, \alpha_{12}, \gamma_1) = (\alpha_{21}, \alpha_{22}, \gamma_2)$ when $\beta = 1.5$

Cenários	N	Nível nominal 10%						Nível nominal 5%						Nível nominal 1%					
		KL	R0.2	R0.5	R0.9	R1.5	RV	KL	R0.2	R0.5	R0.9	R1.5	RV	KL	R0.2	R0.5	R0.9	R1.5	RV
1	9	0,135	0,282	0,254	0,279	0,261	0,206	0,084	0,221	0,205	0,220	0,194	0,173	0,039	0,154	0,143	0,151	0,124	0,118
	25	0,093	0,175	0,171	0,162	0,218	0,157	0,056	0,124	0,119	0,113	0,157	0,112	0,017	0,066	0,062	0,055	0,083	0,057
	49	0,095	0,137	0,137	0,130	0,168	0,120	0,045	0,075	0,075	0,067	0,107	0,090	0,009	0,026	0,031	0,018	0,031	0,033
	81	0,084	0,110	0,109	0,107	0,128	0,129	0,040	0,059	0,060	0,055	0,072	0,081	0,007	0,017	0,019	0,013	0,020	0,038
	121	0,095	0,098	0,096	0,099	0,115	0,131	0,043	0,049	0,046	0,048	0,060	0,101	0,008	0,012	0,013	0,011	0,015	0,047
	225	0,102	0,102	0,101	0,103	0,112	0,103	0,051	0,052	0,050	0,053	0,062	0,073	0,012	0,014	0,013	0,013	0,018	0,030
2	9	0,144	0,295	0,266	0,290	0,266	0,211	0,091	0,236	0,210	0,222	0,209	0,174	0,032	0,161	0,140	0,147	0,131	0,125
	25	0,096	0,180	0,179	0,167	0,217	0,160	0,048	0,121	0,117	0,112	0,148	0,124	0,011	0,072	0,070	0,057	0,077	0,061
	49	0,074	0,129	0,127	0,121	0,148	0,120	0,038	0,080	0,079	0,073	0,092	0,073	0,008	0,039	0,042	0,029	0,038	0,043
	81	0,092	0,119	0,118	0,119	0,146	0,130	0,049	0,069	0,070	0,068	0,082	0,093	0,009	0,021	0,022	0,017	0,029	0,043
	121	0,100	0,109	0,108	0,110	0,124	0,138	0,056	0,062	0,061	0,063	0,075	0,090	0,014	0,018	0,019	0,017	0,024	0,045
	225	0,099	0,101	0,100	0,101	0,106	0,096	0,044	0,045	0,045	0,045	0,050	0,068	0,010	0,010	0,011	0,011	0,014	0,031
3	9	0,125	0,193	0,170	0,207	0,243	0,207	0,079	0,123	0,099	0,135	0,183	0,171	0,022	0,045	0,037	0,052	0,109	0,122
	25	0,093	0,141	0,135	0,150	0,230	0,158	0,051	0,082	0,076	0,089	0,167	0,126	0,019	0,025	0,020	0,026	0,088	0,063
	49	0,098	0,127	0,124	0,131	0,177	0,129	0,057	0,073	0,070	0,076	0,114	0,090	0,013	0,015	0,013	0,017	0,046	0,046
	81	0,089	0,113	0,111	0,115	0,142	0,125	0,048	0,063	0,061	0,066	0,083	0,087	0,009	0,013	0,012	0,014	0,027	0,037
	121	0,092	0,104	0,103	0,105	0,123	0,117	0,042	0,051	0,049	0,052	0,068	0,085	0,006	0,010	0,010	0,010	0,013	0,035
	225	0,117	0,117	0,117	0,117	0,124	0,126	0,065	0,065	0,065	0,065	0,071	0,087	0,013	0,013	0,013	0,013	0,018	0,029

Fonte: Elaborada pelo autor (2023)

Tabela 6 – Taxas de Rejeição de Testes de Divergência sob $H0 : (\alpha_{11}, \alpha_{12}, \gamma_1) = (\alpha_{21}, \alpha_{22}, \gamma_2)$ when $\beta = 2$

Cenários	N	Nível nominal 10%						Nível nominal 5%						Nível nominal 1%					
		KL	R0.2	R0.5	R0.9	R1.5	RV	KL	R0.2	R0.5	R0.9	R1.5	RV	KL	R0.2	R0.5	R0.9	R1.5	RV
1	9	0,123	0,294	0,282	0,293	0,256	0,221	0,085	0,229	0,210	0,226	0,190	0,181	0,026	0,165	0,152	0,150	0,106	0,128
	25	0,086	0,197	0,200	0,183	0,215	0,153	0,047	0,141	0,144	0,129	0,150	0,111	0,014	0,090	0,096	0,074	0,081	0,057
	49	0,096	0,159	0,162	0,144	0,167	0,127	0,053	0,101	0,108	0,088	0,113	0,087	0,013	0,053	0,060	0,037	0,047	0,038
	81	0,092	0,123	0,120	0,117	0,131	0,117	0,042	0,070	0,072	0,064	0,077	0,081	0,010	0,033	0,034	0,021	0,024	0,043
	121	0,088	0,103	0,104	0,106	0,123	0,127	0,042	0,056	0,057	0,054	0,066	0,086	0,006	0,017	0,021	0,014	0,016	0,035
	225	0,106	0,105	0,104	0,106	0,113	0,110	0,058	0,058	0,056	0,058	0,066	0,073	0,011	0,011	0,011	0,012	0,013	0,027
2	9	0,147	0,328	0,304	0,317	0,278	0,202	0,094	0,264	0,240	0,252	0,224	0,168	0,027	0,187	0,165	0,172	0,128	0,115
	25	0,095	0,198	0,195	0,183	0,224	0,144	0,056	0,143	0,144	0,126	0,149	0,113	0,015	0,090	0,093	0,071	0,074	0,058
	49	0,089	0,162	0,166	0,154	0,165	0,129	0,045	0,105	0,112	0,093	0,110	0,093	0,011	0,058	0,067	0,044	0,046	0,035
	81	0,084	0,117	0,123	0,115	0,135	0,123	0,040	0,067	0,071	0,064	0,077	0,091	0,009	0,032	0,035	0,022	0,020	0,049
	121	0,089	0,103	0,104	0,103	0,115	0,116	0,043	0,057	0,059	0,056	0,069	0,077	0,007	0,019	0,020	0,016	0,017	0,031
	225	0,101	0,102	0,101	0,101	0,109	0,126	0,051	0,051	0,051	0,051	0,059	0,080	0,012	0,012	0,012	0,012	0,015	0,030
3	9	0,173	0,229	0,204	0,248	0,248	0,202	0,114	0,146	0,119	0,164	0,180	0,161	0,046	0,060	0,050	0,069	0,106	0,108
	25	0,087	0,132	0,125	0,136	0,220	0,162	0,049	0,077	0,071	0,081	0,154	0,125	0,016	0,021	0,017	0,023	0,086	0,073
	49	0,071	0,099	0,097	0,103	0,139	0,130	0,037	0,054	0,051	0,058	0,087	0,093	0,008	0,012	0,011	0,014	0,031	0,040
	81	0,097	0,114	0,112	0,118	0,151	0,137	0,045	0,054	0,052	0,056	0,083	0,097	0,009	0,017	0,015	0,017	0,026	0,045
	121	0,092	0,108	0,105	0,108	0,122	0,118	0,042	0,052	0,051	0,052	0,068	0,085	0,008	0,010	0,009	0,011	0,019	0,041
	225	0,098	0,101	0,099	0,101	0,106	0,107	0,052	0,055	0,054	0,055	0,060	0,075	0,014	0,014	0,014	0,014	0,018	0,035

Fonte: Elaborada pelo autor (2023)

Tabela 7 – Taxas de Rejeição de Testes de Divergência sob $H0 : (\alpha_{11}, \alpha_{12}, \gamma_1) = (\alpha_{21}, \alpha_{22}, \gamma_2)$ when $\beta = 2.5$

Cenários	N	Nível nominal 10%						Nível nominal 5%						Nível nominal 1%					
		KL	R0.2	R0.5	R0.9	R1.5	RV	KL	R0.2	R0.5	R0.9	R1.5	RV	KL	R0.2	R0.5	R0.9	R1.5	RV
1	9	0,124	0,319	0,294	0,312	0,246	0,206	0,083	0,263	0,242	0,250	0,188	0,169	0,030	0,193	0,182	0,173	0,107	0,113
	25	0,083	0,208	0,207	0,191	0,212	0,151	0,042	0,148	0,154	0,133	0,148	0,120	0,012	0,101	0,110	0,082	0,078	0,063
	49	0,084	0,158	0,163	0,145	0,165	0,127	0,047	0,114	0,122	0,097	0,107	0,099	0,011	0,063	0,077	0,043	0,046	0,050
	81	0,092	0,140	0,143	0,136	0,137	0,107	0,044	0,093	0,097	0,082	0,085	0,075	0,008	0,046	0,052	0,032	0,020	0,031
	121	0,079	0,092	0,092	0,094	0,107	0,118	0,035	0,048	0,051	0,046	0,057	0,082	0,007	0,019	0,022	0,015	0,013	0,039
	225	0,105	0,104	0,102	0,105	0,113	0,120	0,053	0,050	0,048	0,052	0,059	0,090	0,013	0,012	0,012	0,013	0,015	0,040
2	9	0,140	0,340	0,324	0,337	0,270	0,201	0,091	0,279	0,261	0,271	0,207	0,173	0,039	0,210	0,195	0,196	0,129	0,110
	25	0,082	0,220	0,220	0,191	0,205	0,145	0,047	0,168	0,171	0,143	0,134	0,105	0,012	0,126	0,127	0,091	0,065	0,055
	49	0,085	0,158	0,158	0,147	0,159	0,131	0,045	0,112	0,115	0,099	0,100	0,089	0,008	0,070	0,078	0,051	0,034	0,043
	81	0,083	0,121	0,123	0,116	0,134	0,113	0,039	0,077	0,080	0,066	0,078	0,081	0,012	0,044	0,049	0,028	0,030	0,031
	121	0,085	0,097	0,096	0,096	0,110	0,123	0,041	0,054	0,054	0,050	0,052	0,086	0,006	0,019	0,019	0,014	0,012	0,040
	225	0,098	0,098	0,097	0,099	0,106	0,133	0,047	0,046	0,046	0,046	0,056	0,097	0,008	0,009	0,009	0,009	0,013	0,036
3	9	0,127	0,213	0,186	0,228	0,231	0,190	0,075	0,146	0,125	0,156	0,176	0,162	0,030	0,063	0,055	0,072	0,099	0,127
	25	0,077	0,143	0,134	0,152	0,211	0,153	0,036	0,074	0,068	0,078	0,151	0,116	0,010	0,022	0,020	0,023	0,078	0,067
	49	0,074	0,111	0,107	0,114	0,155	0,119	0,036	0,064	0,059	0,067	0,102	0,090	0,008	0,014	0,013	0,016	0,039	0,043
	81	0,089	0,111	0,109	0,114	0,141	0,099	0,039	0,055	0,054	0,056	0,079	0,072	0,008	0,012	0,012	0,012	0,022	0,031
	121	0,090	0,111	0,108	0,113	0,129	0,116	0,044	0,052	0,052	0,054	0,070	0,081	0,007	0,010	0,011	0,010	0,018	0,030
	225	0,109	0,113	0,112	0,114	0,123	0,114	0,051	0,056	0,055	0,056	0,068	0,075	0,011	0,013	0,012	0,013	0,018	0,035

Fonte: Elaborada pelo autor (2023)

4.7 APLICAÇÃO EM DADOS REAIS

Nesta seção, realiza-se uma aplicação a dados reais extraídos de imagens PolSAR obtidas do sensor de radar de abertura sintética para veículos aéreos não habitados (UAVSAR). O objetivo é resolver problemas de detecção de mudanças em área urbana densa cujas mudanças são causadas pelo processo de urbanização. Registramos duas cenas capturadas pelo sensor UAVSAR referentes a duas áreas distintas da cidade de Los Angeles (CA, EUA).

Para cada região, há duas imagens extraídas de momentos diferentes. O primeiro par - visto nas figuras 15a e 15b- foi tomada em 23 de abril de 2009 e 3 de maio de 2015, respectivamente. O segundo par mostrado nas figuras 17a e 17b- foi obtido em 23 de abril de 2009 e 3 de maio de 2015, respectivamente. A seguir, os detectores de mudança propostos são usados para verificar mudanças na região que ocorreram entre os dois períodos diferentes. Para obter os resultados, utilizamos uma janela de tamanho 3×3 na vizinhança de pixels e os testes baseados em divergência sob o nível de significância 1×10^{-4} .

As figuras 14 — 15 exibem os mapas preditivos para aplicação da Cena 1. Pode-se observar na Tabela 8 que o detector baseado na divergência de Rényi gerou valores do Coeficiente Kappa (CV) próximo de 1, sendo eles: o método $CD_R^{0.9}$ com 60,33% para o canal HH-HV e 52,96% para o canal HH-VV. Já para o canal HV-VV, destacou-se o método $CD_R^{0.2}$ com CV de 63,01%. Quando o CV se aproxima de 1, significa haver uma concordância entre os detectores e as cenas reais da imagem PolSAR, o que indica nesse caso que o detector é relativamente preciso quanto às mudanças em áreas urbanas. Os mapas preditivos podem ser vistos em 15g, 16d e 16i, respectivamente.

Pode-se observar, na Tabela 9, que o detector baseado na divergência de Rényi e o detector baseado na estatística RV geraram valores do CV próximo de 1, sendo eles: o método $CD_R^{0.9}$ com 63,7% para o canal HH-HV e 61,36% para o canal HV-VV. Já para o canal HH-VV foi o método baseado na RV com CV de 60,11%. Os mapas preditivos podem ser vistos em 16 — 17. Visivelmente percebe-se pelo mapa de referência 10c maiores mudanças entre 2009 e 2015 nas regiões localizadas na parte superior da imagem. Observa-se que os detectores $CD_R^{0.9}$ e CD_{RV} produziram mapas preditivos com ruído speckle baixo para todos os canais da Cena 2, o que indica, neste caso, que os detectores são relativamente precisos quanto às mudanças em áreas urbanas.

Tabela 8 – Desempenho dos detectores - Cena 1 - Los Angeles (CA, USA).

Detectores	FP (%)	FN (%)	FA (%)	DR (%)	κ (%)
HH-HV					
S_{KL}	7.519	8.335	109.482	90.253	37.455
$S_R^{0.2}$	4.773	5.574	71.372	93.481	59.201
$S_R^{0.5}$	4.131	7.984	105.864	90.984	47.872
$S_R^{0.9}$	4.887	5.284	68.254	93.791	60.332
S_{RV}	7.416	10.469	232.616	88.659	-6.595
HH-VV					
S_{KL}	7.543	9.753	132.147	88.779	29.063
$S_R^{0.2}$	4.141	7.484	97.248	91.499	50.731
$S_R^{0.5}$	4.131	7.984	105.864	90.984	47.872
$S_R^{0.9}$	4.211	7.052	90.413	91.945	52.962
S_{RV}	7.704	10.451	227.102	88.641	-6.68
HV-VV					
S_{KL}	5.617	7.148	92.721	91.711	48.638
$S_R^{0.2}$	4.428	5.105	65.201	94.021	62.536
$S_R^{0.5}$	4.118	5.498	69.088	93.613	61.472
$S_R^{0.9}$	4.746	4.821	62.836	94.314	63.014
S_{RV}	7.762	10.512	228.674	88.575	-7.391

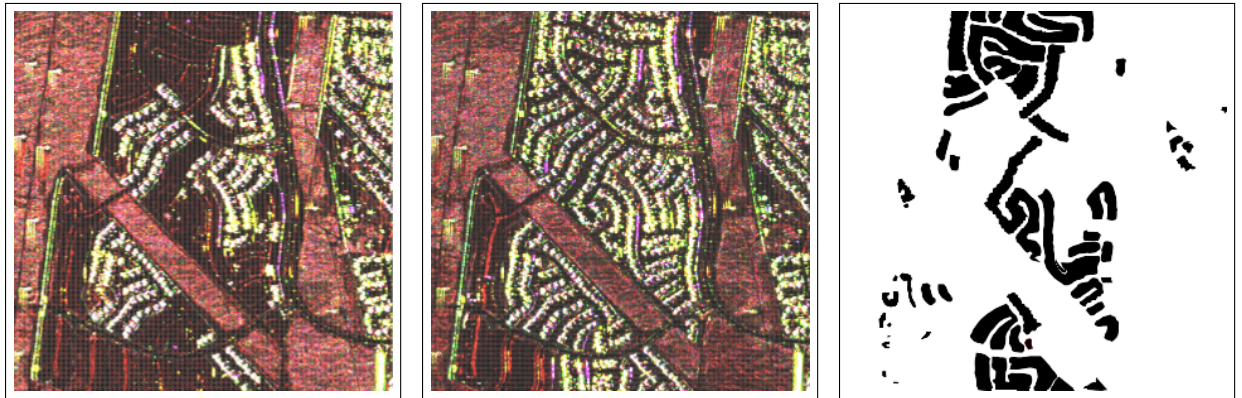
Fonte: Elaborada pelo autor (2023)

Tabela 9 – Desempenho dos detectores - Cena 2 - Los Angeles (CA, USA).

Detectores	FP (%)	FN (%)	FA (%)	DR (%)	κ (%)
HH-HV					
S_{KL}	3.711	5.437	101.469	94.024	48.666
$S_R^{0.2}$	1.819	4.871	86.969	94.725	60.141
$S_R^{0.5}$	2.041	5.294	97.937	94.278	55.875
$S_R^{0.9}$	1.682	4.524	78.566	95.088	63.372
S_{RV}	0.76	5.239	95.805	94.41	61.683
HH-VV					
S_{KL}	3.662	5.543	103.886	93.918	47.991
$S_R^{0.2}$	1.555	6.649	145.261	92.953	45.941
$S_R^{0.5}$	1.799	7.118	164.441	92.474	40.551
$S_R^{0.9}$	1.427	6.172	126.726	93.434	50.814
S_{RV}	0.658	5.48	103.692	94.175	60.116
HV-VV					
S_{KL}	1.443	5.958	118.862	93.646	52.652
$S_R^{0.2}$	0.904	5.525	105.051	94.114	58.646
$S_R^{0.5}$	0.811	5.894	118.494	93.753	55.847
$S_R^{0.9}$	0.991	5.154	93.422	94.483	61.369
S_{RV}	0.529	5.652	110.02	94.011	59.209

Fonte: Elaborada pelo autor (2023)

Figura 14 – Mapas de detecção de mudanças do SPAN HH-HV para região de Los Angeles (CA, USA).



(a) Cena 1 - 2009.

(b) Cena 1 - 2015.

(c) Mapa de referência.



(d) S_{KL} - HH-HV.



(e) $S_R^{0.2}$ - HH-HV.



(f) $S_R^{0.5}$ - HH-HV.



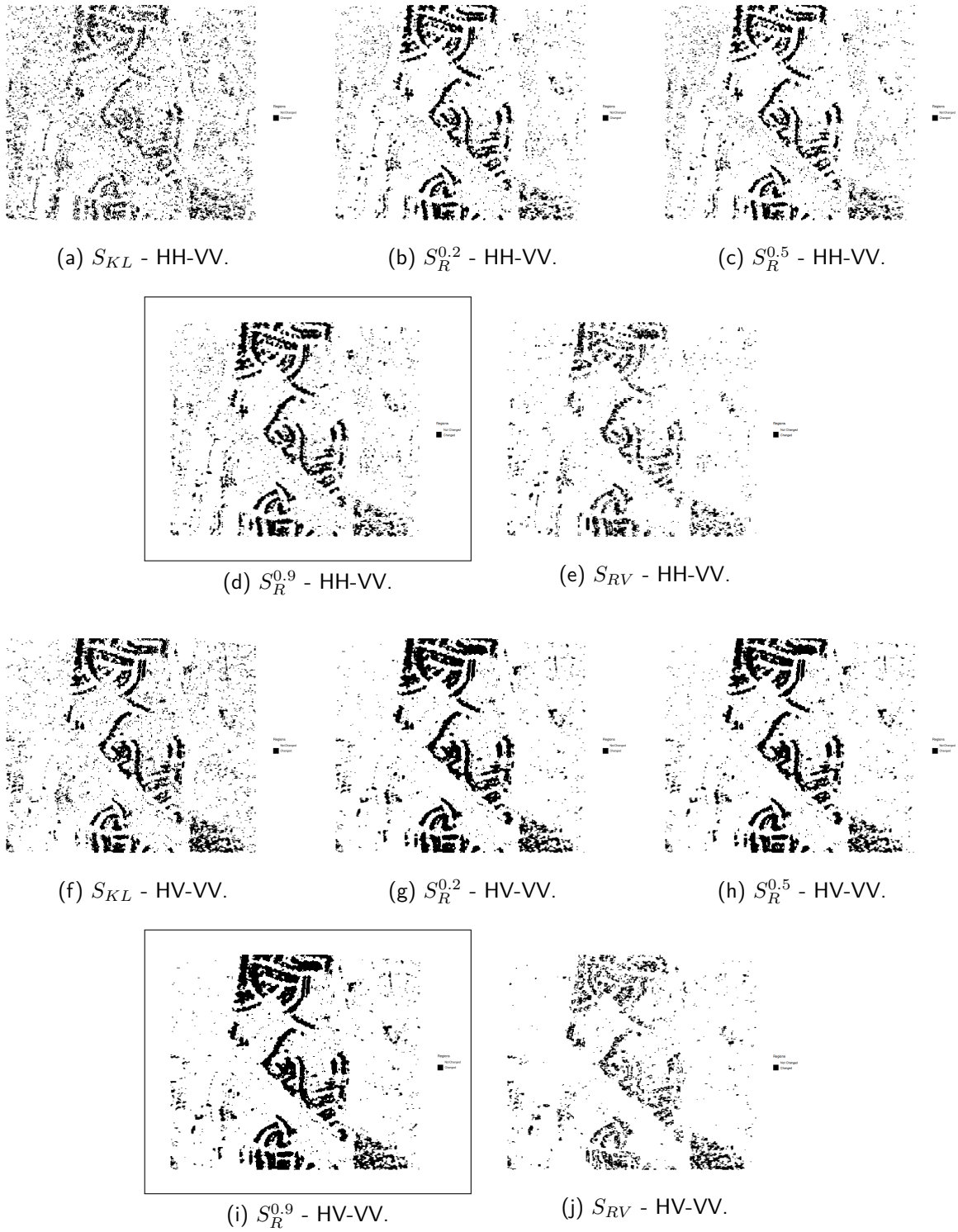
(g) $S_R^{0.9}$ - HH-HV.



(h) S_{RV} - HH-HV.

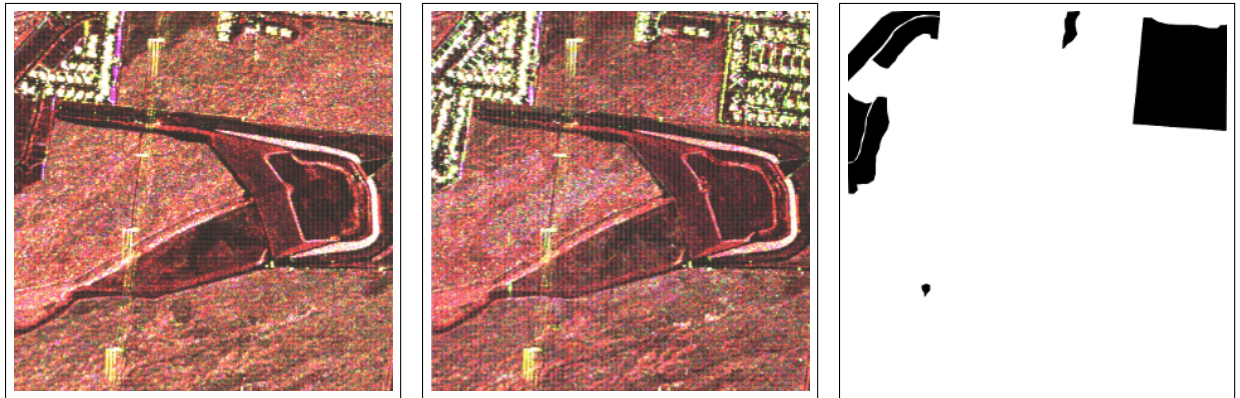
Fonte: Elaborada pelo autor (2023)

Figura 15 – Mapas de detecção de mudanças do SPAN HH-VV e HV-VV para região de Los Angeles (CA, USA).



Fonte: Elaborada pelo autor (2023)

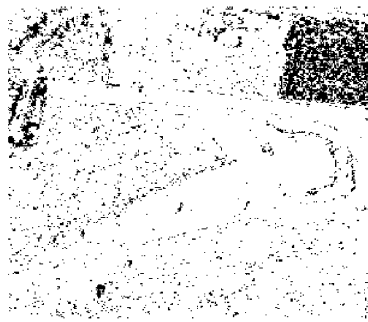
Figura 16 – Mapas de detecção de mudanças do SPAN HH-HV para região de Los Angeles (CA, USA).



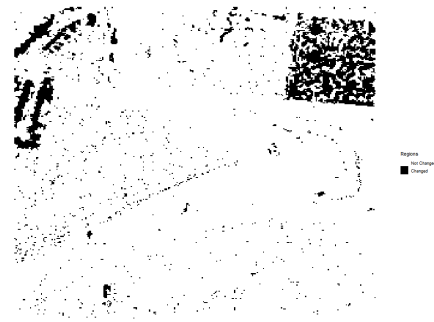
(a) Cena 2 - 2009.

(b) Cena 2 - 2015.

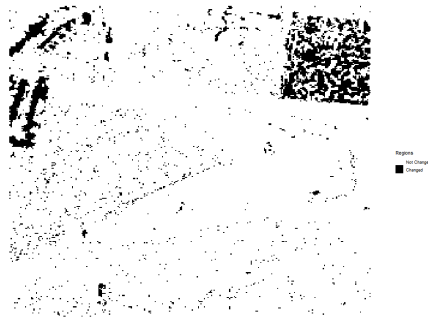
(c) Mapa de referência.



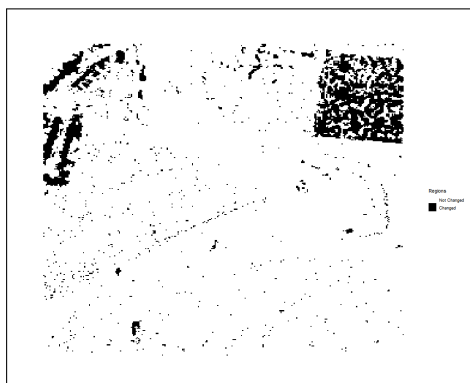
(d) S_{KL} - HH-HV.



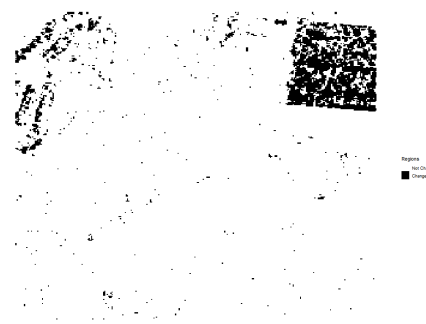
(e) $S_R^{0.2}$ - HH-HV.



(f) $S_R^{0.5}$ - HH-HV.



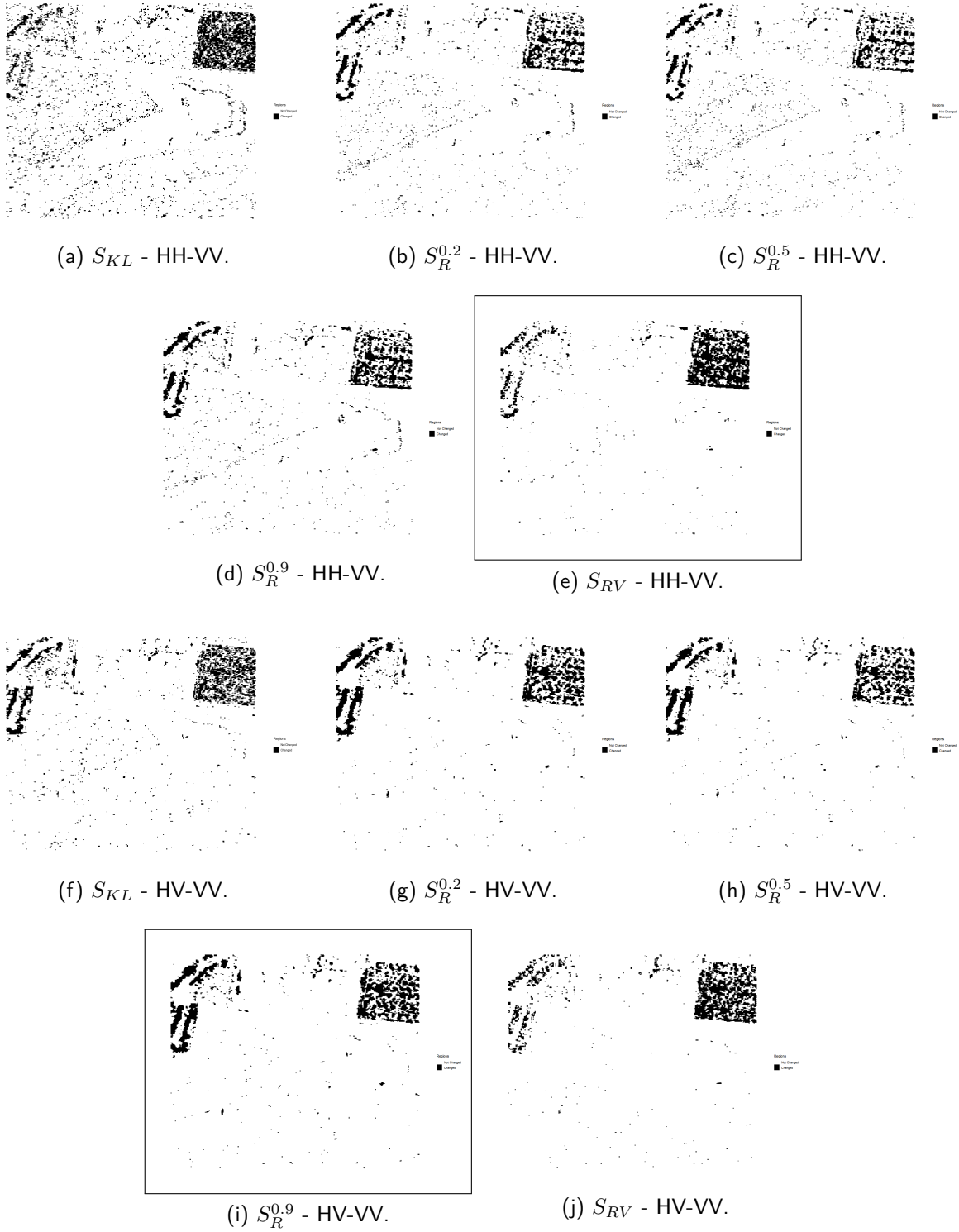
(g) $S_R^{0.9}$ - HH-HV.



(h) S_{RV} - HH-HV-.

Fonte: Elaborada pelo autor (2023)

Figura 17 – Mapas de detecção de mudanças do SPAN HH-VV e HV-VV para região de Los Angeles (CA, USA).



Fonte: Elaborada pelo autor (2023)

5 CONCLUSÃO

Em virtude de suas vantagens, as imagens de Radar de Abertura Sintética Polarimétrica (PolSAR) foram utilizadas na resolução de problemas de sensoriamento remoto. Vários atributos podem ser estudados a partir do dado PolSAR. Um dos mais importantes deles é o Total Scattered Power, conhecido como "SPAN". Livros e artigos Gao (2019), Lee e Pottier (2009), West e Riley (2019), Ansari, Buddhiraju e Malhotra (2020), Canisius et al. (2018), Wang et al. (2018), Wang et al. (2017), Xu et al. (2014) indicam que essa quantidade carrega significativa informação sobre a imagem PolSAR e suas propriedades estatísticas associadas. A partir da revisão da literatura coberta nesta tese, o SPAN sempre é analisado isoladamente.

Esta tese apresenta uma forma de analisá-lo conjuntamente. De uma perspectiva estatística, esta tese apresentou uma medida probabilística para descrever o par (Intensidade, SPAN) em imagens PolSAR multilook, por meio da lei bivariada Gama de McKay (MBG) ainda não utilizada no contexto aplicado em tela. Como extensão desse suposto, propôs-se também uma nova distribuição bivariada resultante do produto entre as distribuições MBG e Gama Inversa, chamada distribuição $\mathcal{G}_I^0\text{MB}$. Várias propriedades matemáticas em forma fechada foram propostas para a distribuição $\mathcal{G}_I^0\text{MB}$: matriz de covariância, densidades das distribuições condicionais, esperança e variâncias condicionais e funções geradoras de momentos condicionais. As distribuições MBG e $\mathcal{G}_I^0\text{MB}$ descrevem o par aleatório induzido do SPAN e uma das intensidades dos canais de polarização.

Na literatura de dados PolSAR multilook, as distribuições usadas para descrever dados de regiões homogêneas é a Wishart Complexa escalonada. A lei MBG pode ser derivada a partir da diagonal principal da Wishart Complexa escalonada, motivo pela qual essa lei foi escolhida como a primeira suposição. Assim, a distribuição MBG é indicada para pares de regiões homogêneas. Por outro lado, a dedução da $\mathcal{G}_I^0\text{MB}$ no contexto de modelagem multiplicativa sugere que essa distribuição seja usada também para cenários mais heterogêneos. Como discutido na tese, as imagens PolSAR são corrompidas por um ruído multidimensional conhecido por speckle. O speckle afeta multiplicativamente alguns atributos das imagens PolSAR (tais como as intensidades e o SPAN), portanto, os dados resultantes tanto carecem de propostas adequadas na modelagem como nas etapas de pós-processamento.

Uma vez definidas as suposições probabilísticas, esta tese seguiu uma tendência promissora de usar medidas de divergência na elaboração de novos testes de hipótese e estatísticas de

testes para duas amostras que pode ser entendido como o funcional sobre qual se soluciona o problema de detecção de mudança. Nesse ponto, o problema se concentrou em deduzir medidas de divergências analítica e computacionalmente simples e propor o teste de hipótese. Para esse fim, derivaram-se expressões em forma fechada para as divergências e distâncias de Kullback-Leibler (KL) e Rényi para as distribuições MBG e $\mathcal{G}_I^0\text{MB}$. Adicionalmente, propuseram testes de hipótese de duas amostras. A computação dos testes baseados em divergências depende de estimadores com certas propriedades assintóticas, presente no método de máxima verossimilhança. Assim, apresentamos discussões sobre a estimação por máxima verossimilhança para os parâmetros da MBG e $\mathcal{G}_I^0\text{MB}$.

O desempenho dos testes foi quantificado por experimentos de Monte Carlo, adotando como critério de comparação estimativa para níveis nominais usuais na literatura (1%, 5% e 10%). Além disso, os testes propostos são comparados com o teste da razão entre verossimilhanças. Como resultados principais observam-se: para distribuição MBG, o método derivado da divergência de Rényi com ordem 0,5 e, para a nova distribuição bivariada $\mathcal{G}_I^0\text{M}$, o método derivado da divergência de Kullback-Leibler que apresentou melhores resultados. Além disso, os testes de hipóteses baseado em divergências obtiveram melhor desempenho quanto à aproximação das propriedades assintóticas em comparação com os testes da razão entre verossimilhanças para duas amostras. Uma vez avaliado por simulação, os testes de hipóteses também foram aplicados para construir detectores de mudanças em imagens PolSAR das regiões de Los Angeles (Califórnia - EUA). Os resultados dessa aplicação apontaram que:

- Os detectores construídos através dos testes de hipótese baseado na divergência de Rényi de alta ordem para a distribuição MBG são melhores para capturar mudanças centrais na imagem. A partir de um mapa preditivo binário obtido do teste de Rényi da lei MBG, eventos de urbanização no centro da imagem são melhores inferidos. Por outro lado, os detectores construídos através dos testes de hipóteses baseado na divergência KL para a distribuição MBG capturaram mudanças nas regiões mais próximas às bordas da imagem.
- Os detectores construídos através dos testes de hipótese baseado na divergência de Rényi de ordem 0.9 para a distribuição $\mathcal{G}_I^0\text{MB}$ capturaram melhores mudanças do processo de urbanização nas imagens. Vale ressaltar que no par de canais HH-VV, o detector construído através do teste da razão entre verossimilhanças obteve um desempenho melhor do que os demais.

Como comentário final, esta tese apresentou avanços nas grandes áreas de Análise Multivariada, Teoria da Informação e Processamento de dados PolSAR. Embora as contribuições aplicadas tenham se concentrado na detecção de mudanças, elas podem ser aplicadas em várias áreas de pós-processamento de imagens obtidas pelo uso de iluminação coerente (como no caso das imagens PolSAR e médicas), tais como filtragem, segmentação e classificação. Esse fato torna as contribuições interessantes para os usuários potenciais em áreas do processamento de imagens.

5.1 FUTUROS TRABALHOS

- ✓ Propor a distribuição $\mathcal{G}_I^0\text{MB}$ k-variada como extensão ao trabalho do Mathai e Moschopoulos (1991) e derivar alguns de seus resultados teóricos, tais como o vetor de médias, matriz de covariância, funções geradoras de momentos, funções geradoras de cumulantes, densidades condicionais, esperanças e variâncias condicionais;
- ✓ Propor divergências, distâncias, testes de hipóteses e matriz de informação de Fisher para o caso $\mathcal{G}_I^0\text{MB}$ k-variado;
- ✓ Fazer a detecção de mudanças em imagens PolSAR a partir do atributo (HH, HH+HV, SPAN), baseando-se nas divergências de Kullback-Leibler e Rényi para o caso $\mathcal{G}_I^0\text{MB}$ k-variado;
- ✓ Como uma extensão do trabalho de Rahayu, Prastyo et al. (2020), propor modelos de regressão para a distribuição $\mathcal{G}_I^0\text{MB}$ e o seu caso k-variado.

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)*, v. 28, p. 131–142, 1966.
- AMARI, S. *Differential-geometrical Methods in Statistics*. New York-USA: Springer-Verlag, 1985. v. 1. ISBN 9780387960562.
- ANSARI, R. A.; BUDDHIRAJU, K. M.; MALHOTRA, R. Urban change detection analysis utilizing multiresolution texture features from polarimetric SAR images. *Remote Sensing Applications: Society and Environment*, v. 20, p. 100418, 2020.
- BARNES, R. M. Roll-invariant Decompositions for the Polarization Covariance Matrix. In: . Redstone Arsenal, AL: Polarimetry Technology Workshop, 1988.
- BORBA, A. A. de; MARENGONI, M.; FRERY, A. C. Fusion of Evidences for Edge Detection in PolSAR images. In: . Kochi, India: IEEE Recent Advances in Geoscience and Remote Sensing: Technologies, Standards and Applications (TENGARSS), 2019. p. 80–85.
- BOUHLEL, N.; MÉRIC, S. Multilook Polarimetric SAR Change Detection using Stochastic Distances between Matrix-Variate G_d^0 Distributions. *IEEE Transactions on Geoscience and Remote Sensing*, v. 58, p. 6823–6843, 2020.
- CAMERON, W. L.; RAIS, H. Conservative Polarimetric Scatterers and Their Role in Incorrect Extensions of the Cameron Decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, v. 44, p. 3506–3516, 2006.
- CANISIUS, F.; SHANG, J.; LIU, J.; HUANG, X.; MA, B.; JIAO, X.; GENG, X.; KOVACS, J. M.; WALTERS, D. Tracking crop phenological development using multi-temporal polarimetric RADARSAT-2 data. *Remote Sensing of Environment*, v. 210, p. 508–518, 2018.
- CASELLA, G.; BERGER, R. L. *Statistical Inference*. Belmont, CA-USA: Cengage Learning, 2021. v. 2. 1-323 p. ISBN 9780534243128.
- CHAN, Y. K.; KOO, V. An introduction to Synthetic Aperture Radar SAR. *Progress In Electromagnetics Research B*, v. 2, p. 27–60, 2008.
- CLOUDE, S. R.; POTTIER, E. A review of target decomposition theorems in radar polarimetry. *IEEE Transactions on Geoscience and Remote Sensing*, v. 34, p. 498–518, 1996.
- CORDEIRO, G. M.; CINTRA, R. J.; RÊGO, L. C.; NASCIMENTO, A. D. The Gamma generalized Normal distribution: A descriptor of SAR imagery. *Journal of Computational and Applied Mathematics*, v. 347, p. 257–272, 2019.
- COVER, T. M.; THOMAS, J. A. *Elements of Information Theory*. New York: Wiley-Interscience, 2007. 407 p. ISBN 9780471241959.
- CSISZÁR, I. Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*, v. 2, p. 299–318, 1967.
- ERVEN, T. van; HARREMOËS, P. Rényi divergence and majorization. In: . Austin, TX-USA: IEEE International Symposium on Information Theory, 2010. p. 1335–1339.

- FANG, S.; ZHU, Q. Independent Gaussian distributions minimize the Kullback-Leibler (KL) Divergence from Independent Gaussian Distributions. arXiv preprint arXiv:2011.02560, 2020.
- FERREIRA, J. A.; COÊLHO, H.; NASCIMENTO, A. D. C. A family of divergence-based classifiers for Polarimetric Synthetic Aperture Radar (PolSAR) imagery vector and matrix features. *International Journal of Remote Sensing*, v. 42, p. 1201–1229, 2021.
- FISHER, R. A. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, v. 222, p. 309–368, 1922.
- FREEMAN, A.; DURDEN, S. L. A three-component scattering model for polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 36, p. 963–973, 1998.
- 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.; 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 Distributions. *IEEE Transactions on Geoscience and Remote Sensing*, v. 52, p. 1213–1226, 2014.
- GAO, G. *Characterization of SAR Clutter and Its Applications to Land and Ocean Observations*. Beijing, China: Springer, 2019. 1-174 p. ISBN 9789811310195.
- GAO, G.; QIN, X.; ZHOU, S. Modeling SAR images based on a generalized Gamma distribution for texture component. *EMW Publishing - Progress In Electromagnetics Research*, v. 137, p. 669–685, 2013.
- GIRALDEZ, A. E. SAOCOM-1 Argentina L-band SAR mission overview. *ESA Special Publication*, v. 565, p. 27, 2004.
- GUPTA, A. K.; NADARAJAH, S. Intensity-duration models based on bivariate gamma distributions. *Hiroshima Mathematical Journal*, v. 36, p. 387 – 395, 2006.
- HACHICHA, S.; CHAABANE, F. Application of DSM theory for SAR image change detection. In: . Cairo, Egypt: IEEE International Conference on Image Processing (ICIP), 2009. p. 3733–3736.
- HAGEDORN, M.; SMITH, P. J.; BONES, P. J.; MILLANE, R. P.; PAIRMAN, D. A Trivariate Chi-Squared Distribution Derived from the Complex Wishart Distribution. *Journal of Multivariate Analysis*, v. 97, p. 655–674, 2006.
- HAJNSEK, I.; DESNOS, Y.-L. *Polarimetric Synthetic Aperture Radar: Principles and Application*. Cham, Switzerland: Springer Nature, 2020. v. 25. 294 p. ISBN 9783030565022.
- HARTLEY, R. V. L. Transmission of Information. *Bell System Technical Journal*, v. 7, p. 535–563, 1928.
- HORTA, M. M. *Modelos de Mistura de Distribuições na Segmentação de Imagens SAR Polarimétricas Multi-look*. Tese (Doutorado) — Universidade de São Paulo, 2009.

HORTA, M. M.; MASCARENHAS, N. D.; SPORTOUCHE, H.; SEICHEPINE TITLE=Change detection in multitemporal HR SAR images: A hypothesis test-based approach, N.; TUPIN, F.; NICOLAS, J.-M. In: . Munich, Germany: IEEE, 2012. p. 374–377.

HUYNEN, J. Theory and applications of the N-target decomposition theorem. *Journées Internationales de la polarimétrie Radar*, p. 20–22, 1990.

HUYNEN, J. R. *Phenomenological Theory of Radar Targets*. Tese (Doutorado) — Proefschrift—Technische Hogeschool, Delft, Rotterdam, 1970.

INCE, T.; KIRANYAZ, S.; GABBOUJ, M. Evolutionary RBF classifier for polarimetric SAR images. *Expert Systems with Applications*, v. 39, p. 4710–4717, 2012.

KAILATH, T. The Divergence and Bhattacharyya Distance Measures in Signal Selection. *IEEE Transactions on Communication Technology*, v. 15, p. 52–60, 1967.

KROGAGER, E. New decomposition of the radar target scattering matrix. *Electronics Letters*, v. 18, p. 1525–1527, 1990.

KULLBACK, S. *Information Theory and Statistics*. USA: Dover Publications, 1968. 279–335 p. (A Wiley publication in mathematical statistics). ISBN 9780486696843.

KULLBACK, S.; LEIBLER, R. A. On Information and Sufficiency. *The Annals of Mathematical Statistics*, v. 22, p. 79–86, 03 1951.

LAMPORT, L. *Latex, A Document Preparation System*. Boston, MA-USA: Addison-Wesley, Reading, MA, 1986. v. 1. 258 p. ISBN 9780201157901.

LEE, J. S.; HOPPEL, K. W.; MANGO, S. A.; MILLER, A. R. Intensity and Phase Statistics of Multilook Polarimetric and Interferometric SAR Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, v. 32, p. 1017–1028, 1994.

LEE, J.-S.; POTTIER, E. *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton: CRC, 2009. 1–391 p.

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.

LIESE, F.; VAJDA, I. *Convex Statistical Distances*. Leipzig, Germany: Teubner-Texte zur Mathematik, 1987. v. 95. ISBN 9783322004284.

LIMA, W. V. F.; NASCIMENTO, A. D. C.; AMARAL, G. J. A. Distance-based tests for planar shape. *Journal of Multivariate Analysis*, v. 184, p. 104743, 2021. ISSN 0047-259X.

LIVINGSTONE, C.; GRAY, A.; HAWKINS, R.; OLSEN, R. CCRS C/X-airborne Synthetic Aperture Radar: An R and D tool for the ERS-1 time frame. *IEEE Aerospace and Electronic Systems Magazine*, v. 3, p. 11–20, 1988.

MATHAI, A.; MOSCHOPOULOS, P. On a multivariate gamma. *Journal of Multivariate Analysis*, v. 39, p. 135–153, 1991. ISSN 0047-259X.

NADARAJAH, S. Reliability for some bivariate gamma distributions. *Mathematical Problems in Engineering*, p. 843–924, 2005. ISSN 1024-123X.

NADARAJAH, S.; GUPTA, A. K. Some bivariate Gamma distributions. *Applied Mathematics Letters*, v. 19, p. 767–774, 2006.

NASCIMENTO, A. D.; FRERY, A. C.; CINTRA, R. J. Detecting changes in fully polarimetric SAR imagery with statistical information theory. *IEEE Transactions on Geoscience and Remote Sensing*, v. 57, p. 1380–1392, 2018.

POTTIER, E. On Dr. J.R. Huynen's main contributions in the development of polarimetric radar technique. In: . San Diego, CA, United States: SPIE, 1992. v. 1748, p. 72–85.

PRUDNIKOV, A. P.; BRYCHKOV, Y. A.; MARICHEV, O. I.; ROMER, R. H. *Integrals and series: Elementary functions*. King's Lynn-Great Britain: Taylor & Francis, 1986. v. 1. 328 p. ISBN 2881240895.

RAHAYU, A.; PRASTYO, D. D. et al. Multivariate gamma regression: Parameter estimation, hypothesis testing, and its application. *Symmetry-Multidisciplinary Digital Publishing Institute*, v. 12, 2020.

RODRÍGUEZ, W. D. A.; AMARAL, G. J. A.; NASCIMENTO, A. D. C.; FERREIRA, J. A. Information criteria in classification: new divergence-based classifiers. *Journal of Statistical Computation and Simulation*, v. 90, p. 3200–3217, 2020.

SALICRÚ, M.; MORALES, D.; MENÉNDEZ, M.; PARDO, L. On the applications of divergence type measures in testing statistical hypotheses. *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.

SKRIVER, H.; DALL, J.; TOAN, T. L.; QUEGAN, S.; FERRO-FAMIL, L.; POTTIER, E. Agriculture Classification using PolSAR Data. In: LACOSTE, H. (Ed.). Frascati, Italy: the 2nd International Workshop POLINSAR 2005 (ESA SP-586), 2005.

THOMPSON, T. A user's guide for the NASA/JPL Synthetic Aperture Radar and the NASA/JPL L-and C-band scatterometers. JPL publication, p. 38–83, 1986.

VAN, Z. Application of Cloude's target decomposition theorem to polarimetric imaging radar. In: . San Diego, CA-USA: SPIE., 1992. v. 127, p. 184–211.

WALKER, B.; SANDER, G.; THOMPSON, M.; BURNS, B.; FELLERHOFF, R.; DUBBERT, D. A high-resolution, four-band SAR Testbed with real-time image formation. In: . Lincoln, NE-USA: IGARSS'96. 1996 International Geoscience and Remote Sensing Symposium, 1996. v. 3, p. 1881–1885.

WANG, H.; HAN, J.; DENG, Y. PolSAR image classification based on Laplacian Eigenmaps and superpixels. *EURASIP Journal on Wireless Communications and Networking*, p. 198, 2017.

WANG, S.; HUA, W.; LIU, H.; JIAO, L. Unsupervised classification for polarimetric SAR images based on the improved CFSFDP algorithm. *International Journal of Remote Sensing*, v. 40, p. 3154–3178, 2019.

- WANG, W.; XIANG, D.; BAN, Y.; ZHANG, J.; WAN, J. Superpixel Segmentation of Polarimetric SAR Images Based on Integrated Distance Measure and Entropy Rate Method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 10, p. 4045–4058, 2017.
- WANG, W.; XIANG, D.; BAN, Y.; ZHANG, J.; WAN, J. Enhanced edge detection for polarimetric SAR images using a directional span-driven adaptive window. *International Journal of Remote Sensing*, v. 39, p. 6340–6357, 2018.
- WASHAYA, P.; BALZ, T.; MOHAMADI, B. Coherence change-detection with sentinel-1 for natural and anthropogenic disaster monitoring in urban areas. *Remote Sensing*, v. 10, p. 1026, 2018.
- WEST, R. D.; RILEY, R. M. Polarimetric Interferometric SAR Change Detection Discrimination. *IEEE Transactions on Geoscience and Remote Sensing*, United States, v. 57, p. 3091–3104, 2019.
- XU, L.; LI, S.; DENG, Y.; WANG, R. Unsupervised classification of polarimetric synthetic aperture radar interferometry using polarimetric interferometric similarity parameters and SPAN. *IET Radar, Sonar & Navigation*, v. 8, p. 1135–1144, 2014.
- YAMAGUCHI, Y.; MORIYAMA, T.; ISHIDO, M.; YAMADA, H. Four-component scattering model for polarimetric SAR image decomposition. *IEEE Transactions on geoscience and remote sensing*, v. 43, p. 1699–1706, 2005.
- YAMAGUCHI, Y.; YAJIMA, Y.; YAMADA, H. A four-component decomposition of POLSAR images based on the coherency matrix. *IEEE Geoscience and Remote Sensing Letters*, v. 3, p. 292–296, 2006.
- YANG, J.; PENG, Y. N.; YAMAGUCHI, Y.; YAMADA, H. On Huynen's decomposition of a Kennaugh matrix. *IEEE Geoscience and Remote Sensing Letters*, v. 3, p. 369–372, 2006.
- YANG, J.; YAMAGUCHI, Y.; YAMADA, H.; SENGOKU, M.; LIN, S. M. Stable decomposition of a Kennaugh matrix,. *IEICE Transaction Communications*, E81-B, p. 1261–1268, 1998.
- YANG, R.; DAI, B.; TAN, L.; LIU, X.; YANG, Z.; LI, H. *Polarimetric Microwave Imaging*. Singapore: Springer Nature, 2021. v. 1. 591 p. ISBN 97789811588969.
- YUE, S.; OUARDA, T.; BOBÉE, B. A review of bivariate gamma distributions for hydrological application. *Journal of Hydrology*, v. 246, p. 1–18, 2001. ISSN 0022-1694.
- ZHANG, L.; SUN, L.; ZOU, B.; MOON, W. M. Fully Polarimetric SAR Image Classification via Sparse Representation and Polarimetric Features. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 8, p. 3923–3932, 2015.
- ZYL, J. J. V. *Synthetic Aperture Radar Polarimetry*. Hoboken, NJ-USA: John Wiley & Sons, 2011. v. 2. 327 p. ISBN 9781118115114.

APÊNDICE A – DIVERGÊNCIA DE KULLBACK-LEIBLER PARA DISTRIBUIÇÃO MCKAY

A divergência Kullback-Leibler entre dois pares $\mathbf{x} \sim \text{MBG}(\alpha_{11}, \alpha_{12}, \gamma_1)$ e $\mathbf{y} \sim \text{MBG}(\alpha_{21}, \alpha_{22}, \gamma_2)$ é considerado

$$\begin{aligned} D_{KL}(\mathbf{x}||\mathbf{y}) &= \int_0^\infty \int_0^{x_2} f(x_1, x_2; \alpha_{11}, \alpha_{12}, \gamma_1) \log \frac{f(x_1, x_2; \alpha_{11}, \alpha_{12}, \gamma_1)}{f(x_1, x_2; \alpha_{21}, \alpha_{22}, \gamma_2)} dx_1 dx_2 \\ &= \int_0^\infty \int_0^{x_2} \frac{x_1^{\alpha_{11}-1} (x_2 - x_1)^{\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right)}{\gamma_1^{\alpha_{11}+\alpha_{12}} \Gamma(\alpha_{11}) \Gamma(\alpha_{12})} \\ &\quad \times \log \left(\underbrace{\frac{x_1^{\alpha_{11}-1} (x_2 - x_1)^{\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right)}{\gamma_1^{\alpha_{11}+\alpha_{12}} \Gamma(\alpha_{11}) \Gamma(\alpha_{12})} \frac{\gamma_2^{\alpha_{21}+\alpha_{22}} \Gamma(\alpha_{21}) \Gamma(\alpha_{22})}{x_1^{\alpha_{21}-1} (x_2 - x_1)^{\alpha_{22}-1} \exp\left(-\frac{x_2}{\gamma_2}\right)}}_A \right) dx_1 dx_2 \end{aligned}$$

O termo A pode ser reescrito como

$$A = \underbrace{\log \left(\frac{\gamma_2^{\alpha_{21}+\alpha_{22}} \Gamma(\alpha_{21}) \Gamma(\alpha_{22})}{\gamma_1^{\alpha_{11}+\alpha_{12}} \Gamma(\alpha_{11}) \Gamma(\alpha_{12})} \right)}_{A1} + \underbrace{\log \left(x_1^{\alpha_{11}-\alpha_{21}} (x_2 - x_1)^{\alpha_{12}-\alpha_{22}} \exp \left(-\frac{x_2}{\gamma_1} + \frac{x_2}{\gamma_2} \right) \right)}_{A2}.$$

Então,

$$\begin{aligned} D_{KL}(\mathbf{x}||\mathbf{y}) &= \frac{1}{\gamma_1^{\alpha_{11}+\alpha_{12}} \Gamma(\alpha_{11}) \Gamma(\alpha_{12})} A1 \overbrace{\int_0^\infty \int_0^{x_2} x_1^{\alpha_{11}-1} (x_2 - x_1)^{\alpha_{12}-1} \exp \left(-\frac{x_2}{\gamma_1} \right) dx_1 dx_2}^{W_1} \\ &\quad + \frac{1}{\gamma_1^{\alpha_{11}+\alpha_{12}} \Gamma(\alpha_{11}) \Gamma(\alpha_{12})} \underbrace{\int_0^\infty \int_0^{x_2} x_1^{\alpha_{11}-1} (x_2 - x_1)^{\alpha_{12}-1} \exp \left(-\frac{x_2}{\gamma_1} \right) A2 dx_1 dx_2}_{W_2}. \end{aligned}$$

Obtemos o seguinte resultado para o primeiro período

$$W_1 = \int_0^\infty \int_0^{x_2} x_1^{\alpha_{11}-1} (x_2 - x_1)^{\alpha_{12}-1} \exp \left(-\frac{x_2}{\gamma_1} \right) dx_1 dx_2.$$

Note que $(x_2 - x_1)^{\alpha_{12}-1} = \left[x_2 \cdot \left(1 - \frac{x_1}{x_2} \right) \right]^{\alpha_{12}-1} = x_2^{\alpha_{12}-1} \cdot \left(1 - \frac{x_1}{x_2} \right)^{\alpha_{12}-1}$.

Assumindo x_2 constante e fazendo a variável mudar $t = \frac{x_1}{x_2} \in (0, 1)$, então $dx_1 = x_2 dt$.

Portanto, nós temos,

$$\begin{aligned}
W_1 &= \int_0^\infty x_2^{\alpha_{11}+\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) \int_0^1 t^{\alpha_{11}-1} \cdot (1-t)^{\alpha_{12}-1} dt dx_2 \\
&= \text{Beta}(\alpha_{11}, \alpha_{12}) \int_0^\infty x_2^{\alpha_{11}+\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) dx_2.
\end{aligned}$$

Portanto,

$$\begin{aligned}
W_1 &= \text{Beta}(\alpha_{11}, \alpha_{12}) \cdot \int_0^\infty x_2^{\alpha_{11}+\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) dx_2 \\
&= \text{Beta}(\alpha_{11}, \alpha_{12}) \cdot \Gamma(\alpha_{11} + \alpha_{12}) \cdot \gamma_1^{\alpha_{11}+\alpha_{12}} = \gamma_1^{\alpha_{11}+\alpha_{12}} \cdot \Gamma(\alpha_{11}) \cdot \Gamma(\alpha_{12}).
\end{aligned}$$

Para o segundo termo, temos:

$$\begin{aligned}
W_2 &= \underbrace{\int_0^\infty \int_0^{x_2} x_1^{\alpha_{11}-1} (x_2 - x_1)^{\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) (\alpha_{11} - \alpha_{21}) \log x_1 dx_1 dx_2}_{W_{2A}} \\
&\quad + \underbrace{\int_0^\infty \int_0^{x_2} x_1^{\alpha_{11}-1} (x_2 - x_1)^{\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) (\alpha_{12} - \alpha_{22}) \log(x_2 - x_1) dx_1 dx_2}_{W_{2B}} \\
&\quad - \underbrace{\int_0^\infty \int_0^{x_2} x_1^{\alpha_{11}-1} (x_2 - x_1)^{\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) \left(x_2 \frac{\gamma_2 - \gamma_1}{\gamma_1 \gamma_2}\right) dx_1 dx_2}_{W_{2C}}
\end{aligned}$$

Considere a possibilidade de derivar W_{2A} . Assumindo x_2 constante e fazendo com que a variável mude $t = \frac{x_1}{x_2} \in (0, 1)$, então $dx_1 = x_2 dt$. Assim, nós temos,

$$\begin{aligned}
W_{2A} &= (\alpha_{11} - \alpha_{21}) \int_0^\infty x_2^{\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) \int_0^{x_2} x_1^{\alpha_{11}-1} \log x_1 \left(1 - \frac{x_1}{x_2}\right)^{\alpha_{12}-1} dx_1 dx_2 \\
&= (\alpha_{11} - \alpha_{21}) \int_0^\infty x_2^{\alpha_{11}+\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) \left[\int_0^1 t^{\alpha_{11}-1} (1-t)^{\alpha_{12}-1} \log t dt \right. \\
&\quad \left. + \log x_2 \int_0^1 t^{\alpha_{11}-1} (1-t)^{\alpha_{12}-1} dt \right] dx_2 \\
&= (\alpha_{11} - \alpha_{21}) \text{Beta}(\alpha_{11}, \alpha_{12}) \Gamma(\alpha_{11} + \alpha_{12}) \gamma_1^{\alpha_{11}+\alpha_{12}} [\Psi(\alpha_{11}) + \log \gamma_1].
\end{aligned}$$

Considere derivar $W2_B$ e aplicar os mesmos argumentos anteriores, tem-se

$$\begin{aligned} W2_B &= (\alpha_{12} - \alpha_{22}) \int_0^\infty x_2^{\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) \log x_2 \int_0^{x_2} x_1^{\alpha_{11}-1} \left(1 - \frac{x_1}{x_2}\right)^{\alpha_{12}-1} dx_1 dx_2 \\ &+ (\alpha_{12} - \alpha_{22}) \int_0^\infty x_2^{\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) \int_0^{x_2} x_1^{\alpha_{11}-1} \left(1 - \frac{x_1}{x_2}\right)^{\alpha_{12}-1} \log\left(1 - \frac{x_1}{x_2}\right) dx_1 dx_2 \\ &= (\alpha_{12} - \alpha_{22}) \text{Beta}(\alpha_{11}, \alpha_{12}) \Gamma(\alpha_{11} + \alpha_{12}) \gamma_1^{\alpha_{11} + \alpha_{12}} [\log \gamma_1 + \Psi(\alpha_{12})]. \end{aligned}$$

Considere derivar $W2_C$ e aplicar os mesmos argumentos anteriores, tem-se

$$\begin{aligned} W2_C &= \left(\frac{\gamma_1 + \gamma_2}{\gamma_1 \gamma_2}\right) \cdot \int_0^\infty \int_0^{x_2} x_1^{\alpha_{11}-1} (x_2 - x_1)^{\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) \cdot x_2 dx_1 dx_2 \\ &= \left(\frac{\gamma_1 + \gamma_2}{\gamma_1 \gamma_2}\right) \text{Beta}(\alpha_{11}, \alpha_{12}) (\alpha_{11} + \alpha_{12}) \Gamma(\alpha_{11} + \alpha_{12}) \gamma_1^{\alpha_{11} + \alpha_{12}}. \end{aligned}$$

Assim, após algumas manipulações algébricas, tem-se

$$\begin{aligned} D_{KL}(\mathbf{x}||\mathbf{y}) &= D_{KL}(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2) = \log \frac{\gamma_2^{\alpha_{21} + \alpha_{22}} \Gamma(\alpha_{21}) \Gamma(\alpha_{22})}{\gamma_1^{\alpha_{11} + \alpha_{12}} \Gamma(\alpha_{11}) \Gamma(\alpha_{12})} + \left\{ (\alpha_{11} - \alpha_{21}) [\Psi(\alpha_{11}) + \log \gamma_1] \right. \\ &\quad \left. + (\alpha_{12} - \alpha_{22}) [\log \gamma_1 + \Psi(\alpha_{12})] - \left(\frac{\gamma_1 + \gamma_2}{\gamma_1 \gamma_2}\right) (\alpha_{11} + \alpha_{12}) \right\}. \end{aligned}$$

APÊNDICE B – DIVERGÊNCIA DE RÉNYI DE ORDEM R PARA DISTRIBUIÇÃO MCKAY

A divergência de Rényi divergence de ordem r entre dois pares $\mathbf{x} \sim \text{MBG}(\alpha_{11}, \alpha_{12}, \gamma_1)$ e $\mathbf{y} \sim \text{MBG}(\alpha_{21}, \alpha_{22}, \gamma_2)$ é dada como

$$\begin{aligned}
 D_{R:r}(\mathbf{x}||\mathbf{y}) &= D_{R:r}(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2) \\
 &= \frac{1}{r(r-1)} \int_{\mathcal{Z}} f^r(x_1, x_2; \alpha_{11}, \alpha_{12}, \gamma_1) f^{1-r}(x_1, x_2; \alpha_{11}, \alpha_{12}, \gamma_1) dx_1 dx_2 \\
 &= \left(\frac{1}{\gamma_1^{\alpha_{11}+\alpha_{12}} \Gamma(\alpha_{11}) \Gamma(\alpha_{12})} \right)^r \cdot \left(\frac{1}{\gamma_2^{\alpha_{21}+\alpha_{22}} \Gamma(\alpha_{21}) \Gamma(\alpha_{22})} \right)^{1-r} \\
 &\quad \underbrace{\int_0^\infty \int_0^{x_2} \left(x_1^{\alpha_{11}-1} (x_2 - x_1)^{\alpha_{12}-1} \exp\left(-\frac{x_2}{\gamma_1}\right) \right)^r \left(x_1^{\alpha_{21}-1} (x_2 - x_1)^{\alpha_{22}-1} \exp\left(-\frac{x_2}{\gamma_2}\right) \right)^{1-r} dx_1 dx_2}_{A}.
 \end{aligned} \tag{B.1}$$

Assumindo x_2 constante e fazendo a variável mudar $t = \frac{x_1}{x_2} \in (0, 1)$, então $dx_1 = x_2 dt$.

Assim,

$$\begin{aligned}
 D_{R:r}(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2) &= \left(\frac{1}{\gamma_1^{\alpha_{11}+\alpha_{12}} \Gamma(\alpha_{11}) \Gamma(\alpha_{12})} \right)^r \left(\frac{1}{\gamma_2^{\alpha_{21}+\alpha_{22}} \Gamma(\alpha_{21}) \Gamma(\alpha_{22})} \right)^{1-r} \\
 &\quad \int_0^\infty \int_0^{x_2} x_1^{r\alpha_{11}+\alpha_{21}-r\alpha_{21}-1} (x_2 - x_1)^{r\alpha_{12}+\alpha_{22}-r\alpha_{22}-1} \cdot \exp\left(-\frac{rx_2}{\gamma_1} - \frac{(1-r)x_2}{\gamma_2}\right) dx_1 dx_2 \\
 &= \frac{1}{r(1-r)} \left\{ \log \Gamma(\eta_1) + \log \Gamma(\eta_2) + (1-r) \left[\log \Gamma(\alpha_{21}) + \log \Gamma(\alpha_{22}) \right. \right. \\
 &\quad \left. \left. - (\alpha_{21} + \alpha_{22}) \left[\log \gamma_1 - \log(\gamma_1 - \gamma_1 \cdot r + \gamma_2 \cdot r) \right] \right. \right. \\
 &\quad \left. \left. + r \left[(\alpha_{11} + \alpha_{12}) \cdot (\log \gamma_2 - \log(\gamma_1 - \gamma_1 \cdot r + \gamma_2 \cdot r)) - \log \Gamma(\alpha_{11}) + \log \Gamma(\alpha_{12}) \right] \right\},
 \end{aligned}$$

onde $\eta_1 = \alpha_{21} + r\alpha_{11} - r\alpha_{21}$ e $\eta_2 = \alpha_{22} + r\alpha_{12} - r\alpha_{22}$.

APÊNDICE C – NOVA DISTRIBUIÇÃO BIVARIADA PARA DADOS SPECKLED

Seja $Y \sim \Gamma^{-1}(\tau, \beta)$ e $\mathbf{x} \sim MBG(\alpha_1, \alpha_2, \gamma)$ duas v.a.s independentes. Então, seja $\mathbf{z} := \mathbf{x} \times Y$ cuja função densidade conjunta bivariada $f_z(z_1, z_2)$ é definida:

$$\begin{aligned}
 f_z(z_1, z_2) &= \int_0^\infty \frac{1}{y} f_x\left(\frac{z_1}{y}, \frac{z_2}{y}\right) f_y(y) dy \\
 &= \frac{1}{\gamma^{\alpha_1+\alpha_2} \Gamma(\alpha_1) \Gamma(\alpha_2)} \int_0^\infty \frac{1}{y} \left(\frac{z_1}{y}\right)^{\alpha_1-1} \left(\frac{z_2-z_1}{y}\right)^{\alpha_2-1} \exp\left(-\frac{z_2}{y\gamma}\right) f_y(y) dy \\
 &= \frac{z_1^{\alpha_1-1} (z_2-z_1)^{\alpha_2-1}}{\gamma^{\alpha_1+\alpha_2} \Gamma(\alpha_1) \Gamma(\alpha_2)} \int_0^\infty y^{-(\alpha_1+\alpha_2)+1} \exp\left(-\frac{z_2}{\gamma y}\right) f_y(y) dy \\
 &= \frac{z_1^{\alpha_1-1} (z_2-z_1)^{\alpha_2-1}}{\gamma^{\alpha_1+\alpha_2} \Gamma(\alpha_1) \Gamma(\alpha_2)} \int_0^\infty y^{-(\alpha_1+\alpha_2)+1} \exp\left(-\frac{z_2}{\gamma y}\right) \left(\frac{\beta^\tau}{\Gamma(\tau)} y^{-\tau-1} \exp\left(-\frac{\beta}{y}\right)\right) dy \\
 &= \frac{z_1^{\alpha_1-1} (z_2-z_1)^{\alpha_2-1}}{\gamma^{\alpha_1+\alpha_2} \Gamma(\alpha_1) \Gamma(\alpha_2)} \frac{\beta^\tau}{\Gamma(\tau)} \int_0^\infty y^{-(\tau+\alpha_1+\alpha_2)+1-1} \exp\left(-\frac{z_2}{\gamma} + \beta/x\right) dy \\
 &= \frac{z_1^{\alpha_1-1} (z_2-z_1)^{\alpha_2-1}}{\gamma^{\alpha_1+\alpha_2} \Gamma(\alpha_1) \Gamma(\alpha_2)} \frac{\beta^\tau}{\Gamma(\tau)} \frac{\Gamma(\tau + \alpha_1 + \alpha_2 + 1)}{\left(\frac{z_2}{\gamma} + \beta\right)^{\tau+\alpha_1+\alpha_2+1}}, \tag{C.1}
 \end{aligned}$$

para $0 < z_1 < z_2$ e $\beta, \tau, \gamma, \alpha_1, \alpha_2 > 0$.

APÊNDICE D – FUNÇÕES MARGINAIS E FUNÇÃO DE CORRELAÇÃO PARA $\mathcal{G}_I^0 MB$

Seja $z \sim \mathcal{G}_I^0 MB(\theta)$. A função marginal $f_{Z_1}(z_1)$ é definida a seguir,

$$\begin{aligned} f_{Z_1}(z_1) &= \int_0^\infty \frac{\beta^\tau \Gamma(\tau + \alpha_1 + \alpha_2 + 1) z_1^{\alpha_1-1} (z_2 - z_1)^{\alpha_2-1}}{\gamma^{\alpha_1+\alpha_2} \Gamma(\tau) \Gamma(\alpha_1) \Gamma(\alpha_2) \left(\frac{z_2}{\gamma} + \beta\right)^{\tau+\alpha_1+\alpha_2+1}} dz_2 \\ &= \frac{\beta^\tau \Gamma(\tau + \alpha_1 + \alpha_2 + 1) z_1^{\alpha_1-1} \gamma^{\tau+\alpha_1+\alpha_2+1}}{\gamma^{\alpha_1+\alpha_2} \Gamma(\tau) \Gamma(\alpha_1) \Gamma(\alpha_2)} \underbrace{\int_0^\infty \frac{(z_2 - z_1)^{\alpha_2-1}}{(z_2 + \beta\gamma)^{\tau+\alpha_1+\alpha_2+1}} dz_2}_A \end{aligned}$$

Temos que $(z_2 - z_1)^{\alpha_2-1} = \sum_{n=0}^\infty (-z_1)^n z_2^{\alpha_2-n-1} \binom{\alpha_2-1}{n}$. Portanto, reescrevendo a integral A:

$$A = \int_0^\infty \frac{z_2^{\alpha_2-n-1}}{(z_2 + \beta\gamma)^{\tau+\alpha_1+\alpha_2+1}} dz_2 = (\beta\gamma)^{-(\tau+\alpha_1+n+1)} B(\alpha_2 - n, \tau + \alpha_1 - n + 1) \quad (D.1)$$

Os resultados do cálculo da integral (D.1) foram extraídos da equação 24 da seção 2.2.4 do livro (PRUDNIKOV et al., 1986). Portanto,

$$\begin{aligned} f_{Z_1}(z_1) &= \frac{\beta^\tau \Gamma(\tau + \alpha_1 + \alpha_2 + 1) z_1^{\alpha_1-1} \gamma^{\tau+1}}{\Gamma(\tau) \Gamma(\alpha_1) \Gamma(\alpha_2)} \sum_{n=0}^\infty (-z_1)^n \binom{\alpha_2-1}{n} \\ &\quad (\beta\gamma)^{-(\tau+\alpha_1+n+1)} B(\alpha_2 - n, \tau + \alpha_1 - n + 1) \end{aligned}$$

Em que,

$$\begin{aligned} &\sum_{n=0}^\infty (-z_1)^n \binom{\alpha_2-1}{n} \frac{B(\alpha_2 - n, \tau + \alpha_1 - n + 1)}{(\beta\gamma)^{\tau+\alpha_1+n+1}} = \\ &{}_2F_1\left(\frac{-\alpha_1 - \alpha_2 - \tau}{2}, \frac{-\alpha_1 - \alpha_2 - \tau + 1}{2}; -\alpha_1 - \tau; -\frac{4z_1}{\beta\gamma}\right) \end{aligned}$$

Agora, deriva-se os resultados para a função marginal $f_{Z_2}(z_2)$,

$$\begin{aligned} f_{Z_2}(z_2) &= \int_0^{z_2} \frac{\beta^\tau \Gamma(\tau + \alpha_1 + \alpha_2 + 1) z_1^{\alpha_1-1} (z_2 - z_1)^{\alpha_2-1}}{\gamma^{\alpha_1+\alpha_2} \Gamma(\tau) \Gamma(\alpha_1) \Gamma(\alpha_2) \left(\frac{z_2}{\gamma} + \beta\right)^{\tau+\alpha_1+\alpha_2+1}} dz_1 \\ &= \frac{\beta^\tau \Gamma(\tau + \alpha_1 + \alpha_2 + 1)}{\gamma^{\alpha_1+\alpha_2} \Gamma(\tau) \Gamma(\alpha_1) \Gamma(\alpha_2) \left(\frac{z_2}{\gamma} + \beta\right)^{\tau+\alpha_1+\alpha_2+1}} \underbrace{\int_0^{z_2} z_1^{\alpha_1-1} (z_2 - z_1)^{\alpha_2-1} dz_1}_B \end{aligned}$$

Assume-se z_2 variável constante e fazendo uma mudança de variável chamando $t = \frac{z_1}{z_2} \in (0, 1)$ quando $dz_1 = z_2 dt$. Portanto, temos que

$$z_2^{\alpha_1} \int_0^1 t^{\alpha_1-1} (1-t)^{\alpha_2-1} dt = z_2^{\alpha_1} B(\alpha_1, \alpha_2)$$

Portanto, temos

$$f_{Z_2}(z_2) = \frac{\beta^\tau \Gamma(\tau + \alpha_1 + \alpha_2 + 1) B(\alpha_1, \alpha_2)}{\gamma^{\alpha_1 + \alpha_2} \Gamma(\tau) \Gamma(\alpha_1) \Gamma(\alpha_2)} \frac{z_2^{\alpha_1 + \alpha_2 + 1}}{\left(\frac{z_2}{\gamma} + \beta\right)^{\tau + \alpha_1 + \alpha_2 + 1}}$$

Por fim, derivam-se o coeficiente de correlação entre as variáveis z_1 e z_2 . Deriva-se as variâncias de z_1 e z_2 e a covariância entre z_1 e z_2 da 4.3, portanto,

$$Var(z_1) = \frac{\beta^2 \gamma^2 \alpha_1 [(1 + \alpha_1)(\tau - 1) - \alpha_1(\tau - 2)]}{(\tau - 2)(\tau - 1)^2},$$

$$Var(z_2) = \frac{\beta^2 \gamma^2 \alpha_1 \left[\left(1 + \frac{\alpha_2}{\alpha_1}\right)(1 + \alpha_1 + \alpha_2)(\tau - 1) - \alpha_1(\tau - 2)\left(1 + \frac{\alpha_2}{\alpha_1}\right)^2 \right]}{(\tau - 2)(\tau - 1)^2},$$

$$Cov(z_1, z_2) = \frac{\beta^2 \gamma^2 \alpha_1 (\alpha_1 + \alpha_2)}{(\tau - 2)(\tau - 1)^2},$$

Logo,

$$\sigma_{z_1} \cdot \sigma_{z_2} = \frac{\beta^2 \gamma^2}{(\tau - 2)(\tau - 1)^2} \sqrt{\alpha_1(\tau - \alpha_1 - 1) [(\tau - 1)(\alpha_1 + \alpha_2) + (\alpha_1 + \alpha_2)^2]}$$

Então,

$$\begin{aligned} Cor(z_1, z_2) &= \frac{\frac{\beta^2 \gamma^2 \alpha_1 (\alpha_1 + \alpha_2)}{(\tau - 2)(\tau - 1)^2}}{\frac{\beta^2 \gamma^2}{(\tau - 2)(\tau - 1)^2} \sqrt{\alpha_1(\tau - \alpha_1 - 1) [(\tau - 1)(\alpha_1 + \alpha_2) + (\alpha_1 + \alpha_2)^2]}} \\ &= \sqrt{\frac{\alpha_1(\alpha_1 + \alpha_2)}{(\tau - \alpha_1 - 1)(\alpha_1 + \alpha_2 + \tau - 1)}} \end{aligned}$$

APÊNDICE E – FUNÇÃO GERATRIZ DE MOMENTOS PARA $\mathcal{G}_I^0 MB$

Seja $z \sim \mathcal{G}_I^0 MB(\theta)$. Então, a função geratriz de momentos de z defini-se

$$\begin{aligned} M_z(t) &= \mathbb{E}(\exp(t_1 z_1 + t_2 z_2)) = \int_0^\infty \int_0^{z_2} \exp(t_1 z_1 + t_2 z_2) f(z_1, z_2) dz_1 dz_2 \\ &= \varrho \cdot \int_0^\infty \int_0^{z_2} \exp(t_1 z_1 + t_2 z_2) \frac{z_1^{\alpha_1-1} (z_2 - z_1)^{\alpha_2-1}}{\left(\frac{z_2}{\gamma} + \beta\right)^{\tau+\alpha_1+\alpha_2+1}} dz_1 dz_2 \\ &= \varrho \cdot \int_0^\infty \frac{\exp(t_2 z_2)}{(z_2 + \beta\gamma)^{\tau+\alpha_1+\alpha_2+1}} \int_0^{z_2} z_1^{\alpha_1-1} (z_2 - z_1)^{\alpha_2-1} \exp(t_1 z_1) dz_1 dz_2. \end{aligned}$$

Chamando $\varrho = \frac{\gamma (\beta\gamma)^\tau \Gamma(\tau + \alpha_1 + \alpha_2 + 1)}{\Gamma(\tau) \Gamma(\alpha_1) \Gamma(\alpha_2)}$.

Assumindo z_2 variável constante e fazendo mudança de variável $p = \frac{z_1}{z_2} \in (0, 1)$, quando $dz_1 = z_2 dp$ and $\exp(t_1 z_2 p) = \sum_{n=0}^\infty \frac{(t_1 z_2 p)^n}{n!}$ Portanto, temos,

$$\begin{aligned} \int_0^1 p^{\alpha_1-1} (1-p)^{\alpha_2-1} \exp(t_1 z_2 p) dp &= \sum_{n=0}^\infty \frac{t_1^n z_2^{\alpha_1+\alpha_2+n-1}}{n!} \int_0^1 p^{\alpha_1+n-1} (1-p)^{\alpha_2-1} dp \\ &= \sum_{n=0}^\infty \frac{t_1^n z_2^{\alpha_1+\alpha_2+n-1} B(\alpha_1 + n, \alpha_2)}{n!} \end{aligned}$$

Portanto,

$$\int_0^\infty \frac{z_2^{\alpha_1+\alpha_2+n-1} \exp(t_2 z_2)}{(z_2 + \beta\gamma)^{\tau+\alpha_1+\alpha_2+1}} dz_2$$

Temos que $\exp(t_2 z_2) = \sum_{n=0}^\infty \frac{(t_2 z_2)^n}{n!}$,

$$\sum_{n=0}^\infty \frac{(t_2)^n}{n!} \int_0^\infty \frac{z_2^{\alpha_1+\alpha_2+2n-1}}{(z_2 + \beta\gamma)^{\tau+\alpha_1+\alpha_2+1}} dz_2$$

Portanto,

$$\int_0^\infty \frac{z_2^{\alpha_1+\alpha_2+2n-1}}{(z_2 + \beta\gamma)^{\tau+\alpha_1+\alpha_2+1}} dz_2 = (\beta\gamma)^{2n-\tau-1} B(\alpha_1 + \alpha_2 + 2n, \tau - 2n + 1)$$

Os resultados do cálculo da integral anterior foram extraídos da equação 24 da seção 2.2.24 do livro (PRUDNIKOV et al., 1986).

Temos que

$$\begin{aligned}
& \sum_{n=0}^{\infty} \frac{(t_1)^n}{n!} \frac{(t_2)^n}{n!} (\beta\gamma)^{2n-\tau-1} B(\alpha_1 + \alpha_2 + 2n, \tau - 2n + 1) = \\
& \quad \frac{(\beta\gamma)^{-\tau}}{\beta\gamma} B(\alpha_1, \alpha_2) B(\alpha_1 + \alpha_2, \tau + 1) \cdot \\
& {}_3F_4\left(\alpha_1, \frac{\alpha_1 + \alpha_2}{2}, \frac{\alpha_1 + \alpha_2 + 1}{2}; 1, \alpha_1 + \alpha_2, \frac{1 - \tau}{2}, \frac{-\tau}{2}; t_1 t_2 \gamma^2 \beta^2\right)
\end{aligned}$$

Simplificando os resultados obtemos o resultado conforme o teorema.

APÊNDICE F – ESPERANÇA, VARIÂNCIA E MOMENTOS $\mathcal{G}_I^0 MB$ DO TIPO

$$Z_1 \mid Z_2$$

Seja $z \sim \mathcal{G}_I^0 MB(\theta)$. Deriva-se a esperança condicional do tipo $Z_1 \mid Z_2$:

$$\begin{aligned} \mathbb{E}(Z_1 \mid Z_2) &= \int_0^{z_2} z_1 \cdot f(z_1 \mid z_2) dz_1 \\ &= \int_0^{z_2} z_1 \cdot \frac{f(z_1, z_2)}{f_{Z_2}(z_2)} dz_1 \\ &= \frac{1}{B(\alpha_1, \alpha_2) z_2^{\alpha_1 + \alpha_2 + 1}} \int_0^{z_2} z_1^{\alpha_1 - 1 + 1} \cdot (z_2 - z_1)^{\alpha_2 - 1} dz_1 \\ &= \frac{z_2^{\alpha_2 - 1}}{B(\alpha_1, \alpha_2) z_2^{\alpha_1 + \alpha_2 + 1}} \underbrace{\int_0^{z_2} z_1^{\alpha_1 - 1 + 1} \cdot \left(1 - \frac{z_1}{z_2}\right)^{\alpha_2 - 1} dz_1}_C \end{aligned}$$

em que $f(z_1, z_2)$ é a função densidade conjunta $\mathcal{G}_I^0 MB$ e $f_{Z_2}(z_2)$ é a função marginal.

Assume-se z_2 variável constante e fazendo a mudança de variável $t = \frac{z_1}{z_2} \in (0, 1)$, quando $dz_1 = z_2 dt$. Portanto, temos que,

$$C = z_2^{\alpha_1 + 1} \int_0^1 t^{\alpha_1 - 1 + 1} \cdot (1 - t)^{\alpha_2 - 1} dt = z_2^{\alpha_1 + 1} B(\alpha_1 + 1, \alpha_2)$$

Simplificando os termos da expressão anterior, temos (F.1).

$$\mathbb{E}(Z_1 \mid Z_2) = \frac{B(\alpha_1 + 1, \alpha_2) z_2}{B(\alpha_1, \alpha_2)}$$

De modo similar obtêm-se $\mathbb{E}(Z_1^2 \mid Z_2)$

$$\mathbb{E}(Z_1^2 \mid Z_2) = \frac{B(\alpha_1 + 2, \alpha_2) z_2^2}{B(\alpha_1, \alpha_2)}$$

Agora, deriva-se a variância condicional do tipo $\text{Var}(Z_1 \mid Z_2)$:

$$\begin{aligned} \text{Var}(Z_1 \mid Z_2) &= \mathbb{E}(Z_1^2 \mid Z_2) - [\mathbb{E}(Z_1 \mid Z_2)]^2 \\ &= \frac{B(\alpha_1 + 2, \alpha_2) z_2^2}{B(\alpha_1, \alpha_2)} - \left[\frac{B(\alpha_1 + 1, \alpha_2) z_2}{B(\alpha_1, \alpha_2)} \right]^2 \\ &= \left[\frac{B(\alpha_1, \alpha_2) B(\alpha_1 + 2, \alpha_2) - B(\alpha_1 + 1, \alpha_2)^2}{B(\alpha_1, \alpha_2)^2} \right] z_2^2 \end{aligned}$$

Por fim, deriva-se o momento condicional do tipo $M_{Z_1|Z_2}(t)$:

$$\begin{aligned}
M_{Z_1|Z_2}(t) &= \int_0^{z_2} \exp(z_1 t) \cdot f(z_1 | z_2) dz_1 \\
&= \int_0^{z_2} \exp(z_1 t) \cdot \frac{f(z_1, z_2)}{f_{Z_2}(z_2)} dz_1 \\
&= \frac{1}{B(\alpha_1, \alpha_2) z_2^{\alpha_1 + \alpha_2 + 1}} \int_0^{z_2} z_1^{\alpha_1 - 1} \cdot (z_2 - z_1)^{\alpha_2 - 1} \exp(z_1 t) dz_1 \\
&= \frac{z_2^{\alpha_2 - 1}}{B(\alpha_1, \alpha_2) z_2^{\alpha_1 + \alpha_2 + 1}} \underbrace{\int_0^{z_2} z_1^{\alpha_1 - 1} \cdot \left(1 - \frac{z_1}{z_2}\right)^{\alpha_2 - 1} \exp(z_1 t) dz_1}_D
\end{aligned}$$

Assume-se z_2 variável constante e fazendo a mudança de variável $t = \frac{z_1}{z_2} \in (0, 1)$, quando $dz_1 = z_2 dt$.

$$\exp(z_1 t) = \sum_{n=0}^{\infty} \frac{(z_1 t)^n}{n!}$$

Portanto, temos

$$\begin{aligned}
D &= \sum_{n=0}^{\infty} \frac{(z_2)^{\alpha_1 + n}}{n!} \int_0^1 t^{\alpha_1 + 2n - 1} \cdot (1 - t)^{\alpha_2 - 1} dt \\
&= \sum_{n=0}^{\infty} \frac{(z_2)^{\alpha_1 + n}}{n!} B(\alpha_1 + 2n, \alpha_2)
\end{aligned}$$

Simplificando os termos das expressões anteriores, obtemos $M_{Z_1|Z_2}(t)$. Em que

$$\sum_{n=0}^{\infty} \frac{(z_2)^{\alpha_1 + n}}{n!} B(\alpha_1 + 2n, \alpha_2) = B(\alpha_1, \alpha_2) \cdot {}_2F_2\left(\frac{\alpha_1 + 1}{2}, \frac{\alpha_1}{2}; \frac{\alpha_1 + \alpha_2}{2}, \frac{\alpha_1 + \alpha_2 + 1}{2}; z_2\right)$$

APÊNDICE G – ESPERANÇA, VARIÂNCIA E MOMENTOS $\mathcal{G}_I^0 MB$ DO TIPO

$$Z_2 \mid Z_1$$

Seja $z \sim \mathcal{G}_I^0 MB(\theta)$. Deriva-se a esperança condicional do tipo $Z_2 \mid Z_1$:

$$\begin{aligned} &= \int_0^\infty z_2 \cdot f(z_2 \mid z_1) dz_2 \\ &= \int_0^\infty z_2 \cdot \frac{f(z_1, z_2)}{f_{Z_1}(z_1)} dz_2 \\ &= \frac{(-\beta\gamma)^{\tau+\alpha_1+1}}{B(\alpha_2, \tau + \alpha_1 + 1) {}_2F_1\left(\delta_1, \delta_2; \delta_3; -\frac{4z_1}{\beta\gamma}\right)} \underbrace{\int_0^\infty \frac{z_2 \cdot (z_2 - z_1)^{\alpha_2-1}}{(z_2 + \beta\gamma)^{\tau+\alpha_1+\alpha_2+1}} dz_2}_E \end{aligned}$$

$$\text{em que } \delta_1 = \frac{-\tau - \alpha_1 - \alpha_2}{2}, \delta_2 = \frac{-\tau - \alpha_1 - \alpha_2 + 1}{2} \text{ e } \delta_3 = -\tau - \alpha_1$$

Reescrevendo os termos $(z_2 - z_1)^{\alpha_2-1} = \sum_{n=0}^{\infty} (-z_1)^n z_2^{\alpha_2-n-1} \binom{\alpha_2-1}{n}$, portanto

$$\begin{aligned} E &= \sum_{n=0}^{\infty} (-z_1)^n \binom{\alpha_2-1}{n} \int_0^\infty \frac{z_2^{\alpha_2-n-1}}{(z_2 + \beta\gamma)^{\tau+\alpha_1+\alpha_2+1}} dz_2 \\ &= \sum_{n=0}^{\infty} (-z_1)^n \binom{\alpha_2-1}{n} \frac{B(\alpha_2 - n + 1, \tau + \alpha_1 + n)}{(\beta\gamma)^{\tau+\alpha_1+n}} \\ &= (-\beta\gamma)^{-\tau-\alpha_1} B(\alpha_2 + 1, \tau + \alpha_1) {}_2F_1\left(-\alpha_2, \tau + \alpha_1; -\alpha_2 - 1; \frac{z_1}{\beta\gamma}\right) \end{aligned}$$

Simplificando alguns termos das equações anteriores, obtemos $\mathbb{E}(Z_2 \mid Z_1)$

$$\mathbb{E}(Z_2 \mid Z_1) = \frac{(-\beta\gamma)\alpha_2}{(\tau + \alpha_1)} \frac{{}_2F_1\left(-\alpha_2 - 1, \tau + \alpha_1 - 1; -\alpha_2 - 2; \frac{z_1}{\beta\gamma}\right)}{{}_2F_1\left(\delta_1, \delta_2; \delta_3; \frac{-4z_1}{\beta\gamma}\right)}$$

De modo similar, obtêm-se $\mathbb{E}(Z_2^2 \mid Z_1)$

$$\mathbb{E}(Z_2^2 \mid Z_1) = \frac{\beta^2\gamma^2\alpha_2^2}{(\tau + \alpha_1 - 1)(\tau + \alpha_1)} \frac{{}_2F_1\left(-\alpha_2 - 1, \tau + \alpha_1 - 1; -\alpha_2 - 2; \frac{z_1}{\beta\gamma}\right)}{{}_2F_1\left(\delta_1, \delta_2; \delta_3; \frac{-4z_1}{\beta\gamma}\right)}$$

Agora, deriva-se a variância condicional do tipo $\text{Var}(Z_2 \mid Z_1)$:

$$\begin{aligned}
\text{Var}(Z_2 | Z_1) &= \mathbb{E}(Z_2^2 | Z_1) - [\mathbb{E}(Z_2 | Z_1)]^2 \\
&= \frac{(\beta\gamma\alpha_2)^2 {}_2F_1\left(-\alpha_2 - 1, \tau + \alpha_1 - 1; -\alpha_2 - 2; \frac{z_1}{\beta\gamma}\right)}{(\tau + \alpha_1)(\tau + \alpha_1 - 1) {}_2F_1\left(\delta_1, \delta_2; \delta_3; \frac{-4z_1}{\beta\gamma}\right)} \\
&\quad - \left[\frac{(-\beta\gamma)\alpha_2 {}_2F_1\left(-\alpha_2, \tau + \alpha_1; -\alpha_2 - 1; \frac{z_1}{\beta\gamma}\right)}{\tau + \alpha_1 {}_2F_1\left(\delta_1, \delta_2; \delta_3; \frac{-4z_1}{\beta\gamma}\right)} \right]^2
\end{aligned}$$

Por fim, deriva-se o momento condicional do tipo $M_{Z_2|Z_1}(t)$:

$$\begin{aligned}
M_{Z_2|Z_1}(t) &= \int_0^\infty \exp(z_2 t) \cdot f(z_2 | z_1) dz_2 \\
&= \int_0^\infty \exp(z_2 t) \cdot \frac{f(z_1, z_2)}{f_{Z_1}(z_1)} dz_2 \\
&= \frac{(-\beta\gamma)^{\tau+\alpha_1+1}}{B(\alpha_2, \tau + \alpha_1 + 1) {}_2F_1\left(\delta_1, \delta_2; \delta_3; \frac{-4z_1}{\beta\gamma}\right)} \underbrace{\int_0^\infty \frac{(z_2 - z_1)^{\alpha_2-1} \exp(z_2 t)}{(z_2 + \beta\gamma)^{\tau+\alpha_1+\alpha_2+1}} dz_2}_F
\end{aligned}$$

Reescrevendo os termos $(z_2 - z_1)^{\alpha_2-1} = \sum_{n=0}^\infty (-z_1)^n z_2^{\alpha_2-n-1} \binom{\alpha_2-1}{n}$ e $\exp(z_2 t) = \sum_{n=0}^\infty \frac{(z_2 t)^n}{n!}$, portanto,

$$\begin{aligned}
F &= \sum_{n=0}^\infty (-z_1)^n \binom{\alpha_2-1}{n} \frac{t^n}{n!} \int_0^\infty \frac{(z_2)^{\alpha_2-n-1+n}}{(z_2 + \beta\gamma)^{\tau+\alpha_1+\alpha_2+1}} dz_2 \\
&= \sum_{n=0}^\infty (-z_1)^n \binom{\alpha_2-1}{n} \frac{t^n}{n!} \frac{B(\alpha_2-1, \tau + \alpha_1 + 2)}{(\beta\gamma)^{\tau+\alpha_1+2}}
\end{aligned}$$

Temos que $\sum_{n=0}^\infty (-z_1)^n \binom{\alpha_2-1}{n} \frac{t^n}{n!} = {}_1F_1(-\alpha_1 - 1; 1; -tz_1)$. Simplificando os termos das expressões anteriores obtêm-se $M_{Z_2|Z_1}(t)$.

APÊNDICE H – DIVERGÊNCIA DE KULLBACK-LEIBLER PARA DISTRIBUIÇÃO BIVARIADA $\mathcal{G}_I^0\text{MB}$

A divergência Kullback-Leibler entre dois pares $\mathbf{x} \sim \mathcal{G}_1^0(\alpha_{11}, \alpha_{12}, \gamma_1, \beta_1, \tau_1)$ e $\mathbf{y} \sim \mathcal{G}_2^0(\alpha_{21}, \alpha_{22}, \gamma_2, \beta_2, \tau_2)$ é considerado

$$\begin{aligned} D_{KL}(\mathbf{x}||\mathbf{y}) &= \int_0^\infty \int_0^{x_2} f(z_1, z_2; \alpha_{11}, \alpha_{12}, \gamma_1, \beta_1, \tau_1) \log \frac{f(z_1, z_2; \alpha_{11}, \alpha_{12}, \gamma_1, \beta_1, \tau_1)}{f(z_1, z_2; \alpha_{21}, \alpha_{22}, \gamma_2, \beta_2, \tau_2)} dz_1 dz_2 \\ &= \int_0^\infty \int_0^{z_2} \underbrace{\frac{\beta_1^{\tau_1} \cdot \Gamma((\tau_1 + \alpha_{11} + \alpha_{12} + 1))}{\gamma_1^{\alpha_{11} + \alpha_{12}} \cdot \Gamma(\alpha_{11}) \cdot \Gamma(\alpha_{12}) \cdot \Gamma(\tau_1)}}_A \cdot \frac{z_1^{\alpha_{11}-1} \cdot (z_2 - z_1)^{\alpha_{12}-1}}{\left(\frac{z_2}{\gamma_1} + \beta_1\right)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} \\ &\quad \log \left\{ \frac{\beta_1^{\tau_1} \cdot \Gamma(\tau_1 + \alpha_{11} + \alpha_{12} + 1)}{\gamma_1^{\alpha_{11} + \alpha_{12}} \cdot \Gamma(\alpha_{11}) \cdot \Gamma(\alpha_{12}) \cdot \Gamma(\tau_1)} \cdot \frac{z_1^{\alpha_{11}-1} \cdot (z_2 - z_1)^{\alpha_{12}-1}}{\left(\frac{z_2}{\gamma_1} + \beta_1\right)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} \cdot \right. \\ &\quad \left. \frac{\beta_2^{\tau_2} \cdot \Gamma(\tau_2 + \alpha_{21} + \alpha_{22} + 1)}{\gamma_2^{\alpha_{21} + \alpha_{22}} \cdot \Gamma(\alpha_{21}) \cdot \Gamma(\alpha_{22}) \cdot \Gamma(\tau_2)} \cdot \frac{z_1^{\alpha_{21}-1} \cdot (z_2 - z_1)^{\alpha_{22}-1}}{\left(\frac{z_2}{\gamma_2} + \beta_2\right)^{\tau_2 + \alpha_{21} + \alpha_{22} + 1}} \right\} dz_1 dz_2 \end{aligned}$$

O segundo termo pode ser reescrito como:

$$\begin{aligned} &\underbrace{\log \left(\frac{\beta_1^{\tau_1} \cdot \Gamma(\tau_1 + \alpha_{11} + \alpha_{12} + 1)}{\gamma_1^{\alpha_{11} + \alpha_{12}} \cdot \Gamma(\alpha_{11}) \cdot \Gamma(\alpha_{12}) \cdot \Gamma(\tau_1)} \cdot \frac{\gamma_2^{\alpha_{21} + \alpha_{22}} \cdot \Gamma(\alpha_{21}) \cdot \Gamma(\alpha_{22}) \cdot \Gamma(\tau_2)}{\beta_2^{\tau_2} \cdot \Gamma(\tau_2 + \alpha_{21} + \alpha_{22} + 1)} \right)}_{B1} + \\ &\quad \underbrace{\log \left(\frac{z_1^{\alpha_{11}-1} \cdot (z_2 - z_1)^{\alpha_{12}-1}}{\left(\frac{z_2}{\gamma_1} + \beta_1\right)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} \cdot \frac{\left(\frac{z_2}{\gamma_2} + \beta_2\right)^{\tau_2 + \alpha_{21} + \alpha_{22} + 1}}{z_1^{\alpha_{21}-1} \cdot (z_2 - z_1)^{\alpha_{22}-1}} \right)}_{B2} \end{aligned}$$

Assim,

$$\begin{aligned} D_{KL}(\mathbf{x}||\mathbf{y}) &= A \cdot B_1 \cdot \gamma_1^{\tau_1 + \alpha_{11} + \alpha_{12} + 1} \cdot \underbrace{\int_0^\infty \int_0^{z_2} \frac{z_1^{\alpha_{11}-1} \cdot (z_2 - z_1)^{\alpha_{12}-1}}{\left(z_2 + \beta_1 \gamma_1\right)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} dz_1 dz_2}_{W1} \\ &\quad + A \cdot \gamma_1^{\tau_1 + \alpha_{11} + \alpha_{12} + 1} \cdot \underbrace{\int_0^\infty \int_0^{z_2} \frac{z_1^{\alpha_{11}-1} \cdot (z_2 - z_1)^{\alpha_{12}-1} B_2}{\left(z_2 + \beta_1 \gamma_1\right)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} dz_1 dz_2}_{W2} \end{aligned}$$

Obtemos o seguinte resultado para o primeiro termo,

$$W1 = \int_0^\infty \int_0^{z_2} z_1^{\alpha_{11}-1} (z_2 - z_1)^{\alpha_{12}-1} dz_1 dz_2$$

Temos que $(z_2 - z_1)^{\alpha_{12}-1} = \left[z_2 \cdot \left(1 - \frac{z_1}{z_2} \right) \right]^{\alpha_{12}-1} = z_2^{\alpha_{12}-1} \cdot \left(1 - \frac{z_1}{z_2} \right)^{\alpha_{12}-1}$ Assumindo z_2 constante e fazendo a variável mudar $t \cdot z_2 \rightarrow t = \frac{z_1}{z_2} \in (0, 1)$, então $dz_1 = z_2 dt$. Portanto, nós temos,

$$\begin{aligned} W1 &= \int_0^{z_2} z_1^{\alpha_{11}-1} z_2^{\alpha_{12}-1} \left(1 - \frac{z_1}{z_2} \right)^{\alpha_{12}-1} dz_1 \\ &= z_2^{\alpha_{11}} B(\alpha_{11}, \alpha_{12}) \end{aligned}$$

Portanto,

$$\begin{aligned} W1 &= B(\alpha_{11}, \alpha_{12}) \int_0^\infty \frac{z_2^{\alpha_{11}+\alpha_{12}-1}}{(z_2 + \beta_1 \gamma_1)^{\tau_1+\alpha_{11}+\alpha_{12}+1}} dz_2 \\ &= \gamma_1^{\tau_1+\alpha_{11}+\alpha_{12}+1} \cdot B(\alpha_{11}, \alpha_{12}) \cdot B(\alpha_{11} + \alpha_{12}, \tau_1 + 1) \cdot (\beta_1 \gamma_1)^{-\tau_1-1} \end{aligned}$$

Portanto, ao segundo termo dividiremos desta forma:

$$\begin{aligned} W2 &= (\alpha_{11} - \alpha_{21}) \underbrace{\int_0^\infty \int_0^{z_2} \frac{z_1^{\alpha_{11}} (z_2 - z_1)^{\alpha_{12}-1} \log z_1}{(z_2 + \beta_1 \gamma_1)^{\tau_1+\alpha_{11}+\alpha_{12}+1}} dz_1 dz_2}_{W2_A} + \\ &\quad (\alpha_{12} - \alpha_{22}) \underbrace{\int_0^\infty \int_0^{z_2} \frac{z_1^{\alpha_{11}} (z_2 - z_1)^{\alpha_{12}-1} \log(z_2 - z_1)}{(z_2 + \beta_1 \gamma_1)^{\tau_1+\alpha_{11}+\alpha_{12}+1}} dz_1 dz_2}_{W2_B} + \\ &\quad (\tau_2 + \alpha_{21} + \alpha_{22} + 1) \underbrace{\int_0^\infty \int_0^{z_2} \frac{z_1^{\alpha_{11}} (z_2 - z_1)^{\alpha_{12}-1} (\log(z_2 + \beta_2 \gamma_2) - \log \gamma_2)}{(z_2 + \beta_1 \gamma_1)^{\tau_1+\alpha_{11}+\alpha_{12}+1}} dz_1 dz_2}_{W2_C} + \\ &\quad (\tau_1 + \alpha_{11} + \alpha_{12} + 1) \underbrace{\int_0^\infty \int_0^{z_2} \frac{z_1^{\alpha_{11}} (z_2 - z_1)^{\alpha_{12}-1} (\log \gamma_1 - \log(z_2 + \beta_1 \gamma_1))}{(z_2 + \beta_1 \gamma_1)^{\tau_1+\alpha_{11}+\alpha_{12}+1}} dz_1 dz_2}_{W2_D} \end{aligned}$$

Considere a possibilidade de derivar $W2_A$. Assumindo z_2 constante e fazendo a variável mudar $t = \frac{z_1}{z_2} \in (0, 1)$, então $dz_1 = z_2 dt$. Portanto, nós temos,

$$\begin{aligned}
W2_A &= (\alpha_{11} - \alpha_{21}) \int_0^\infty \frac{1}{(z_2 + \beta_1 \gamma_1)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} \int_0^{z_2} z_1^{\alpha_{11} - 1} (z_2 - z_1)^{\alpha_{12} - 1} \log z_1 dz_1 dz_2 \\
&= (\alpha_{11} - \alpha_{21}) \int_0^\infty \frac{1}{(z_2 + \beta_1 \gamma_1)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} \left\{ \int_0^1 t^{\alpha_{11} - 1} (1 - t)^{\alpha_{12} - 1} \log t dt + \right. \\
&\quad \left. \log z_2 \int_0^1 t^{\alpha_{11} - 1} (1 - t)^{\alpha_{12} - 1} dt \right\} dz_2 \\
&= (\alpha_{11} - \alpha_{21}) [\Psi(\alpha_{11}) + \log(\beta_1 \gamma_1) - \Psi(\tau_1 + 1)] B(\alpha_{11}, \alpha_{12}) \\
&\quad B(\alpha_{11} + \alpha_{12}, \tau_1 + 1) (\beta_1 \gamma_1)^{-\tau_1 - 1}
\end{aligned}$$

Considerar derivar $W2_B$ e aplicar os mesmos argumentos anteriores, tem-se

$$\begin{aligned}
W2_B &= (\alpha_{12} - \alpha_{22}) \int_0^\infty \frac{1}{(z_2 + \beta_1 \gamma_1)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} \int_0^{z_2} z_1^{\alpha_{11} - 1} (z_2 - z_1)^{\alpha_{12} - 1} \log(z_2 - z_1) dz_1 dz_2 \\
&= (\alpha_{12} - \alpha_{22}) \int_0^\infty \frac{1}{(z_2 + \beta_1 \gamma_1)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} \left\{ \log z_2 \int_0^1 t^{\alpha_{11} - 1} (1 - t)^{\alpha_{12} - 1} \right. \\
&\quad \left. + \int_0^1 t^{\alpha_{11} - 1} (1 - t)^{\alpha_{12} - 1} \log(1 - t) dt \right\} dz_2 \\
&= (\alpha_{12} - \alpha_{22}) [\Psi(\alpha_{12}) + \log(\beta_1 \gamma_1) - \Psi(\tau_1 + 1)] B(\alpha_{11}, \alpha_{12}) \\
&\quad B(\alpha_{11} + \alpha_{12}, \tau_1 + 1) (\beta_1 \gamma_1)^{-\tau_1 - 1}
\end{aligned}$$

Considerar derivar $W2_C$ e aplicar os mesmos argumentos anteriores, tem-se

$$\begin{aligned}
W2_C &= (\tau_2 + \alpha_{21} + \alpha_{22} + 1) \int_0^\infty \frac{\log(z_2 + \beta_2 \gamma_2) - \log \gamma_2}{(z_2 + \beta_1 \gamma_1)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} \int_0^{z_2} z_1^{\alpha_{11} - 1} (z_2 - z_1)^{\alpha_{12} - 1} dz_1 dz_2 \\
&= (\tau_2 + \alpha_{21} + \alpha_{22} + 1) \cdot B(\alpha_{11}, \alpha_{12}) \cdot B(\alpha_{11} + \alpha_{12}, \tau_1 + 1) \cdot (\beta_1 \gamma_1)^{-\tau_1 - 1} \\
&\quad \left\{ \log \beta_2 \gamma_2 - \log \gamma_2 + \Psi(\tau_1 + \alpha_{11} + \alpha_{12} + 1) - \Psi(\tau_1 + 1) - \right. \\
&\quad \left. \frac{(\alpha_{11} + \alpha_{12})(1 - \frac{\beta_1 \gamma_1}{\beta_2 \gamma_2})}{\tau_1 + \alpha_{11} + \alpha_{12} + 1} \cdot {}_3F_2 \left(\alpha_{11} + \alpha_{12} + 1, 1, 1; 2, \tau_1 + \alpha_{11} + \alpha_{12} + 2; 1 - \frac{\beta_1 \gamma_1}{\beta_2 \gamma_2} \right) \right\}
\end{aligned}$$

Considerar derivar $W2_D$ e aplicar os mesmos argumentos anteriores, tem-se

$$\begin{aligned}
W2_D &= (\tau_1 + \alpha_{11} + \alpha_{12} + 1) \int_0^\infty \frac{\log \gamma_1 - \log(z_2 + \beta_1 \gamma_1)}{(z_2 + \beta_1 \gamma_1)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} \int_0^{z_2} z_1^{\alpha_{11} - 1} (z_2 - z_1)^{\alpha_{12} - 1} dz_1 dz_2 \\
&= (\tau_1 + \alpha_{11} + \alpha_{12} + 1) B(\alpha_{11}, \alpha_{12}) B(\alpha_{11} + \alpha_{12}, \tau_1 + 1) (\beta_1 \gamma_1)^{-\tau_1 - 1} \\
&\quad [\log \gamma_1 - \log(\beta_1 \gamma_1) - \Psi(\tau_1 + \alpha_{11} + \alpha_{12} + 1) + \Psi(\tau_1 + 1)]
\end{aligned}$$

Assim, após algumas manipulações algébricas, tem-se

$$\begin{aligned}
D_{KL}(\mathbf{x}||\mathbf{y}) &= D_{KL}(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2) = \\
&\frac{\tau_1}{\beta_1} \left\{ \log \left(\frac{\beta_1^{\tau_1} \cdot \Gamma(\tau_1 + \alpha_{11} + \alpha_{12} + 1)}{\gamma_1^{\alpha_{11} + \alpha_{12}} \cdot \Gamma(\alpha_{11}) \cdot \Gamma(\alpha_{12}) \cdot \Gamma(\tau_1)} \cdot \frac{\gamma_2^{\alpha_{21} + \alpha_{22}} \cdot \Gamma(\alpha_{21}) \cdot \Gamma(\alpha_{22}) \cdot \Gamma(\tau_2)}{\beta_2^{\tau_2} \cdot \Gamma(\tau_2 + \alpha_{21} + \alpha_{22} + 1)} \right) + \right. \\
&\quad (\alpha_{11} - \alpha_{21}) [\Psi(\alpha_{11}) + \log(\beta_1 \gamma_1) - \Psi(\tau_1 + 1)] + \\
&\quad (\alpha_{12} - \alpha_{22}) [\Psi(\alpha_{12}) + \log(\beta_1 \gamma_1) - \Psi(\tau_1 + 1)] + \\
&\quad (\tau_2 + \alpha_{21} + \alpha_{22} + 1) \left\{ \log \beta_2 \gamma_2 - \log \gamma_2 + \Psi(\tau_1 + \alpha_{11} + \alpha_{12} + 1) - \Psi(\tau_1 + 1) - \right. \\
&\quad \left. \frac{(\alpha_{11} + \alpha_{12})(1 - \frac{\beta_1 \gamma_1}{\beta_2 \gamma_2})}{\tau_1 + \alpha_{11} + \alpha_{12} + 1} \cdot {}_3F_2 \left(\alpha_{11} + \alpha_{12} + 1, 1, 1; 2, \tau_1 + \alpha_{11} + \alpha_{12} + 2; 1 - \frac{\beta_1 \gamma_1}{\beta_2 \gamma_2} \right) \right\} + \\
&\quad \left. (\tau_1 + \alpha_{11} + \alpha_{12} + 1) [\Psi(\tau_1 + 1) - \Psi(\tau_1 + \alpha_{11} + \alpha_{12} + 1) - \log \beta_1] \right\}.
\end{aligned}$$

APÊNDICE I – DIVERGÊNCIA DE RÉNYI DE ORDEM r PARA DISTRIBUIÇÃO BIVARIADA $\mathcal{G}_I^0\text{MB}$

A divergência de Rényi em ordem de r entre dois pares $\mathbf{x} \sim \mathcal{G}_1^0(\alpha_{11}, \alpha_{12}, \gamma_1, \beta_1, \tau_1)$ e $\mathbf{y} \sim \mathcal{G}_2^0(\alpha_{21}, \alpha_{22}, \gamma_2, \beta_2, \tau_2)$ é considerado

$$\begin{aligned} D_{R:r}(\mathbf{x}||\mathbf{y}) &= D_{R:r}(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2) \\ &= \frac{1}{r(r-1)} \int_{\mathbf{Z}} f^r(z_1, z_2; \alpha_{11}, \alpha_{12}, \gamma_1, \beta_1, \tau_1) f^{1-r}(z_1, z_2; \alpha_{21}, \alpha_{22}, \gamma_2, \beta_2, \tau_2) dz_1 dz_2 \\ &= \left[\frac{\beta_1^{\tau_1} \cdot \Gamma(\tau_1 + \alpha_{11} + \alpha_{12} + 1)}{\gamma_1^{\alpha_{11} + \alpha_{12}} \cdot \Gamma(\alpha_{11}) \cdot \Gamma(\alpha_{12}) \cdot \Gamma(\tau_1)} \right]^r \left[\frac{\beta_2^{\tau_2} \cdot \Gamma(\tau_2 + \alpha_{21} + \alpha_{22} + 1)}{\gamma_2^{\alpha_{21} + \alpha_{22}} \cdot \Gamma(\alpha_{21}) \cdot \Gamma(\alpha_{22}) \cdot \Gamma(\tau_2)} \right]^{1-r} \\ &\quad \underbrace{\int_0^\infty \int_0^{z_2} \left[\frac{z_1^{\alpha_{11}-1} \cdot (z_2 - z_1)^{\alpha_{12}-1}}{\left(\frac{z_2}{\gamma_1} + \beta_1\right)^{\tau_1 + \alpha_{11} + \alpha_{12} + 1}} \right]^r \left[\frac{z_1^{\alpha_{21}-1} \cdot (z_2 - z_1)^{\alpha_{22}-1}}{\left(\frac{z_2}{\gamma_2} + \beta_2\right)^{\tau_2 + \alpha_{21} + \alpha_{22} + 1}} \right]^{1-r} dz_1 dz_2}_A \end{aligned}$$

Assumindo x_2 constante e fazendo a variável mudar $t = \frac{x_1}{x_2} \in (0, 1)$, então $dx_1 = x_2 dt$, assim,

$$\begin{aligned} D_{R:r}(\boldsymbol{\theta}_1||\boldsymbol{\theta}_2) &= \left[\frac{\beta_1^{\tau_1} \cdot \Gamma(\tau_1 + \alpha_{11} + \alpha_{12} + 1)}{\gamma_1^{\alpha_{11} + \alpha_{12}} \cdot \Gamma(\alpha_{11}) \cdot \Gamma(\alpha_{12}) \cdot \Gamma(\tau_1)} \right]^r \left[\frac{\beta_2^{\tau_2} \cdot \Gamma(\tau_2 + \alpha_{21} + \alpha_{22} + 1)}{\gamma_2^{\alpha_{21} + \alpha_{22}} \cdot \Gamma(\alpha_{21}) \cdot \Gamma(\alpha_{22}) \cdot \Gamma(\tau_2)} \right]^{1-r} \\ &\quad \gamma_1^{\tau_1 + \alpha_{11} + \alpha_{12} + 1} \cdot \gamma_2^{\tau_2 + \alpha_{21} + \alpha_{22} + 1} \cdot \int_0^\infty \int_0^{z_2} \frac{z_1^{r(\alpha_{11}-1) + (1-r)(\alpha_{21}-1)} \cdot (z_2 - z_1)^{r(\alpha_{12}-1) + (1-r)(\alpha_{22}-1)}}{(z_2 + \gamma_1\beta_1)^{r(\tau_1 + \alpha_{11} + \alpha_{12} + 1)} \cdot (z_2 + \gamma_2\beta_2)^{(1-r)(\tau_2 + \alpha_{21} + \alpha_{22} + 1)}} dz_1 dz_2 \\ &= \frac{1}{r(1-r)} \left\{ r \left[\log \Gamma(\tau_1 + \alpha_{11} + \alpha_{12} + 1) - \log \Gamma(\tau_1) - \log \Gamma(\alpha_{11}) - \log \Gamma(\alpha_{12}) - \log \beta_1 \right] - \right. \\ &\quad (1-r) \left[\log \Gamma(\tau_2 + \alpha_{21} + \alpha_{22} + 1) - \log \Gamma(\tau_2) - \log \Gamma(\alpha_{21}) - \log \Gamma(\alpha_{22}) - \log \beta_2 - \right. \\ &\quad \left. (\alpha_{21} + \alpha_{22})(\log \beta_2 \gamma_2 - \log \beta_1 \gamma_1) \right] + \log \text{Beta}(\eta_1, \eta_2) + \\ &\quad \log \text{Beta}((1-r)(\tau_2 + \alpha_{21} + \alpha_{22} + 1), r\tau_1 + (1-r)\tau_2 + 1) + \\ &\quad \left. \log {}_3F_2 \left(\eta_1 + \eta_2, (1-r)\delta_2; r\delta_1 + (1-r)\delta_2; 1 - \frac{\beta_1\gamma_1}{\beta_2\gamma_2} \right) \right\}, \end{aligned}$$

em que $\eta_1 = \alpha_{21} + r\alpha_{11} - r\alpha_{21}$, $\eta_2 = \alpha_{22} + r\alpha_{12} - r\alpha_{22}$, $\delta_1 = \tau_1 + \alpha_{11} + \alpha_{12} + 1$ e $\delta_2 = \tau_2 + \alpha_{21} + \alpha_{22} + 1$