Universidade Estadual de Maringá
Departamento de Estatística
Programa de Pós-graduação em Bioestatística

Marta Eliane Echeverria Borges

ANÁLISE DO TAMANHO DAS SEQUÊNCIAS DE DIAS CHUVOSOS VIA DISTRIBUIÇÃO GAMA GENERALIZADA DISCRETA TRUNCADA EM ZERO

Maringá - PR
2020

Marta Eliane Echeverria Borges

ANÁLISE DO TAMANHO DAS SEQUÊNCIAS DE DIAS CHUVOSOS VIA DISTRIBUIÇÃO GAMA GENERALIZADA DISCRETA TRUNCADA EM ZERO

Dissertação apresentada ao Programa de Pós-Graduação em Bioestatística do Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito parcial para a obtenção do título de Mestre em Bioestatística.

Universidade Estadual de Maringá

Departamento de Estatística

Programa de Pós-graduação em Bioestatística

Orientador: Dr. Josmar Mazucheli

Maringá - PR

2020

Dados Internacionais de Catalogação-na-Publicação (CIP)
(Biblioteca Central - UEM, Maringá)

Borges, Marta Eliane Echeverria

B732a Análise do tamanho das sequências de dias chuvosos via distribuição gama generalizada discreta truncada em zero. Marta Eliane Echeverria Borges. Maringá - PR, 2020.

95 f.color, figs., tabs.

Orientador: Prof. Dr. Josmar Mazucheli.

Dissertação (Mestrado) - Universidade Estadual de Maringá, Centro Ciências Exatas, Departamento de Estatística, Programa de Pós-Graduação em Bioestatística, 2020.

1. Distribuições truncadas em zero. 2. Tamanho das sequências de dias chuvosos. 3. Distribuições discretas. 4. Distribuição gama generalizada. I. Mazucheli, Josmar, orient. II. Universidade Estadual de Maringá. Centro de Ciências Exatas. Departamento de Estatística. Programa de Pós-Graduação em Bioestatística. III. Título.

CDD 23.ed. 519.5

Ademir Herinque dos Santos - CRB-9/1065

Marta Eliane Echeverria Borges

ANÁLISE DO TAMANHO DAS SEQUÊNCIAS DE DIAS CHUVOSOS VIA DISTRIBUIÇÃO GAMA GENERALIZADA DISCRETA TRUNCADA EM ZERO

Dissertação apresentada ao Programa de Pós-Graduação em Bioestatística do Centro de Ciências Exatas da Universidade Estadual de Maringá, como requisito parcial para a obtenção do título de Mestre em Bioestatística.

BANCA EXAMINADORA

Prof. Dr. Josmar Mazucheli
Universidade Estadual de Maringá – PBE/UEM

Prof. Dr. Ricardo Puziol de Oliveira
Universidade Estadual de Maringá
Campus Umuarama-PR

Prof. Dr. Eraldo Schunk Silva
Universidade Estadual de Maringá – PBE/UEM

Maringá
21 de maio de 2020

Dedico este trabalho
a minha mãe Wenceslaa e
ao meu pai Valdir.

Agradecimentos

Agradeço a Deus por ter me dado força para superar as dificuldades e por mais um objetivo alcançado.

À Universidade Estadual de Maringá e ao Programa de Pós Graduação em Bioestatística, pela oportunidade de cursar o mestrado.

Ao professor Josmar, pela orientação, tema proposto e pelos conhecimentos compartilhados.

À Agência Nacional de Águas (ANA), ao Instituto Nacional de Meteorologia (INMET) e a ITAIPU Binacional por disponibilizar os dados utilizados neste trabalho.

Aos professores do Programa de Pós Graduação em Bioestatística pelos ensinamentos e conhecimentos transmitidos.

Aos meus familiares e amigos pelo apoio, carinho e compreensão.

À todos que de alguma maneira contribuíram, direta e indiretamente, para a elaboração deste trabalho.

Muito Obrigada!

"O Senhor é a minha força e o meu escudo;
nele o meu coração confia, e dele recebo ajuda."
(Salmos 28:7)

Resumo

Este trabalho teve como objetivo avaliar o desempenho da distribuição gama generalizada discreta truncada em zero para descrever o tamanho das sequências de dias chuvosos das séries históricas anuais das estações meteorológicas das capitais do Brasil. Assim como, também, verificar se os casos particulares com 1 parâmetro da distribuição citada apresentam bons ajustes no tamanho das sequências de dias chuvosos das séries mensais das estações pluviométricas dos Portos São José e Porto Rico do rio Paraná, ou seja, analisar se as distribuições truncadas em zero gama discreta com um parâmetro e a geométrica são boas candidatas para descrever as observações da séries mensais. Para esses mesmos dados a distribuição Lindley discreta truncada em zero também foi avaliada, uma vez que, a mesma já foi recomendada para descrever o tamanhos das sequências de dias chuvosos. Os dados utilizados para estudo foram obtidos a partir dos registros do portal do Instituto Nacional de Meteorologia para maioria das capitais do Brasil, já para as estações de Campo Grande, Porto São José e Porto Rico foram disponibilizados da base de dados do portal da Agência Nacional de Águas. Os resultados indicaram que a distribuição gama generalizada discreta truncada em zero pode ser recomendada na descrição no tamanho das sequências de dias chuvosos para as estações avaliadas, bem como as distribuições truncadas em zero: gama discreta com um parâmetro, geométrica e Lindley discreta que se mostraram adequadas para descrever o tamanho das sequências de dias chuvosos mensais das estações dos Portos avaliados.

Palavras-chave: distribuições truncadas em zero; distribuições discretas, distribuição gama generaliza, tamanho das sequências de dias chuvosos.

Abstract

This work aimed to evaluate the performance of the zero truncated discrete generalized gamma distribution to describe of the length of wet spells of annual historical series of meteorological stations in the capitals of Brazil. As well as to analyze whether the particular cases with 1 parameter of the mentioned distribution have good adjustments of the length of wet spells of the monthly series of the pluviometric stations of the Porto São José and Porto Rico of the Paraná River, that is, to analyze whether the zero truncated distributions discrete gamma with one parameter and geometric are good candidates to describe the observations monthly series. For these same data, the zero truncated discrete Lindley distribution was also evaluated, since it has been recommended to describe of the length of wet spells. The data used for the study were obtained from the records of the portal of the National Meteorological Institute for most capitals of Brazil, whereas for the stations of Campo Grande, Porto São José and Porto Rico were made available from the portal database of the National Water Agency. The results indicated that the zero truncated discrete generalized gamma distribution can be recommended description in the length of wet spells for the evaluated stations, as well as the truncated distributions at zero: discrete gamma with one parameter, geometric and discrete Lindley proved to be adequate to describe describe the length of wet spells monthly at the stations of the evaluated ports.

Keywords: zero truncated distributions; discrete distributions; generalized gamma distribution; length of wet spells.

Lista de ilustrações

Figura 1 – Comportamento da função massa de probabilidade de $GaGD(k, \theta, c)$. . . . 27
Figura 2 – Comportamento da função massa de probabilidade da distribuição $GeoTZ(q)$. 31
Figura 3 – Comportamento da função massa de probabilidade da distribuição $LiDTZ(\beta)$ 32
Figura 4 – Comportamento da funções massa de probabilidade $GaGDTZ(k, \theta, c)$ e $GaGD(k, \theta, c)$. 34
Figura 5 – Representação da contagem do tamanho das sequências dos dias chuvosos. 40
Figura 6 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Nordeste de 1961 a 2018. 47
Figura 7 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Nordeste 1961 a 2018. 48
Figura 8 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Nordeste de 1961 a 2018. 49
Figura 9 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Norte de 1961 a 2018. 55
Figura 10 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Norte de 1961 a 2018. 56
Figura 11 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Centro-Oeste de 1961 a 2018. 60
Figura 12 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Sudeste de 1961 a 2018. 65

Figura 13 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Sul de 1961 a 2018. 70
Figura 14 – Ajuste do tamanho das sequências de dias chuvosos das séries de janeiro a junho da estação Porto São José do rio Paraná - 1985 a 2018. 76
Figura 15 – Ajuste do tamanho das sequências de dias chuvosos das séries de julho a dezembro da estação porto São José do rio Paraná - 1985 a 2018. 77
Figura 16 – Ajuste do tamanho das sequências de dias chuvosos das séries de janeiro a junho da estação Porto Rico do rio Paraná - 1971 a 2018. 84
Figura 17 – Ajuste do tamanho das sequências de dias chuvosos das séries de julho a dezembro da estação Porto Rico do rio Paraná - 1971 a 2018. 85

Lista de tabelas

Tabela 1 – Tipo de precipitação de acordo com a intensidade. 22
Tabela 2 – Descrição das séries históricas das estações meteorológicas estudadas. . . . 38
Tabela 3 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas anuais de 1961 a 2018 das estações meteorológicas das capitais da Região do Nordeste. 45
Tabela 4 – Teste qui-quadrado de aderência para a distribuição $GaGDTZ(k, \theta, c)$ ajustada as séries históricas das estações meteorológicas das capitais da Região Nordeste de 1961 a 2018. 46
Tabela 5 – Estimativa intervalar dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada do tamanho das sequências de dias chuvosos das capitais da Região Nordeste de 1961 a 2018. 51
Tabela 6 – Teste qui-quadrado de aderência e critérios de informação para os modelos ajustados ao tamanho das sequências de dias chuvosos das capitais da Região Nordeste de 1961 a 2018 - limiar 5. 52
Tabela 7 – Análise descritiva do tamanho das sequências de dias chuvosos da série histórica de 1961 a 2018 das estações meteorológicas das capitais da Região do Norte. 53
Tabela 8 – Teste qui-quadrado de aderência para a distribuição $GaGDTZ(k, \theta, c)$ ajustada as séries históricas das estações meteorológicas das capitais da Região Norte de 1961 a 2018. 54
Tabela 9 – Estimação intervalar dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada do tamanho das sequências de dias chuvosos das capitais da Região Norte de 1961 a 2018. 57
Tabela 10 – Teste qui-quadrado de aderência e critérios de informação para os modelos ajustados no tamanho das sequências de dias chuvosos das capitais da Região Norte de 1961 a 2018 - limiar 5. 58

Tabela 11 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas de 1961 a 2018 das estações meteorológicas das capitais da Região Centro-Oeste. 59
Tabela 12 – Teste qui-quadrado de aderência para a distribuição $GaGDTZ(k, \theta, c)$ ajustada ao tamanho das sequências de dias chuvosos das capitais da Região Centro-Oeste de 1961 a 2018. 61
Tabela 13 – Estimativas dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada do tamanho das sequências de dias chuvosos das capitais da Região Centro-Oeste de 1961 a 2018. 62
Tabela 14 – Teste qui-quadrado de aderência e critérios de informação para os modelos ajustados da série histórica das estação das capitais da Região Centro-Oeste de 1961 a 2018 - limiar 5. 63
Tabela 15 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas de 1961 a 2018 das estações meteorológicas das capitais da região Sudeste. 64
Tabela 16 – Teste qui-quadrado de aderência para a distribuição $GaGDTZ(k, \theta, c)$ ajustada ao tamanho das sequências de dias chuvosos das capitais da Região Sudeste de 1961 a 2018. 66
Tabela 17 – Estimativas dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada do tamanho das sequências de dias chuvosos das capitais da Região Sudeste de 1961 a 2018. 67
Tabela 18 – Teste qui-quadrado de aderência e critério de informação para os modelos ajustados no tamanho das sequências de dias chuvosos das capitais da Região Sudeste de 1961 a 2018 - limiar 5. 68
Tabela 19 – Análise descritiva do tamanho das sequências de dias chuvosos da série histórica de 1961 a 2018 das estações meteorológicas das capitais da região Sul. 69
Tabela 20 – Teste qui-quadrado de aderência para a distribuição $GaGDTZ(k, \theta, c)$ ajustada as séries históricas das estações meteorológicas das capitais da região Sul de 1961 a 2018. 71
Tabela 21 – Estimativas dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada do tamanho das sequências de dias chuvosos das capitais da Região Sul de 1961 a 2018. 72
Tabela 22 – Teste qui-quadrado de aderência e critérios de informação para os modelos ajustados das séries históricas das estação das capitais da região Sul de 1961 a 2018 - limiar 5. 73
Tabela 23 – Análise descritiva do tamanho das sequências de dias chuvosos séries históricas anuais de duas estações do Rio Paraná - limiar 5. 74

Tabela 24 – Teste de aderência e estimativas dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada ao tamanho das sequências de dias chuvosos das séries históricas anuais de duas estações do Rio Paraná - limiar 5. 74
Tabela 25 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas mensais da estação do Porto São José do rio Paraná - 1985 a 2018. 75
Tabela 26 – Teste qui-quadrado de aderência para as distribuições ajustadas as séries históricas mensais da estação Porto São José do rio Paraná - 1985 a 2018. 78
Tabela 27 – Valores observados e esperados do tamanho das sequências de dias chuvosos da estação Porto São José do rio Paraná - 1985 a 2018. 80
Tabela 28 – Critérios de informação para as distribuições ajustadas as séries históricas mensais da estação Porto São José do rio Paraná - 1971 a 2018. 81
Tabela 29 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas mensais da estação do Porto Rico do Rio Paraná - 1971 a 2018. 82
Tabela 30 – Teste qui-quadrado de aderência para as distribuições ajustadas as séries históricas mensais da estação Porto Rico do rio Paraná - 1971 a 2018. . . 83
Tabela 31 – Valores observados e esperados do tamanho das sequências de dias chuvosos da estação Porto Rico do rio Paraná - 1971 a 2018. 86
Tabela 32 – Critérios de informação para as distribuições ajustadas as séries históricas mensais da estação Porto Rico do rio Paraná - 1971 a 2018. 87

Lista de abreviaturas e siglas

$GaGDTZ(k, \theta, c)$ Distribuição gama generalizada discreta truncada em zero

$GaDTZ(k, \theta)$ Distribuição gama discreta truncada em zero

$GaDTZ(k)$ Distribuição gama discreta com 1 parâmetro truncada em

$GeoTZ(q)$ Distribuição geométrica truncada em zero

$LiDTZ(\beta)$ Distribuição Lindley discreta truncada em zero

$LNDTZ(\mu\,,\,\sigma)$ Distribuição log-normal discreta truncada em zero

$WeDTZ(q, c)$ Distribuição Weibull discreta truncada em zero

Sumário

1 Introdução . . . 17
1.1 Justificativa . . . 19
1.2 Objetivo . . . 19
1.3 Organização do Trabalho . . . 20
2 Revisão Bibliográfica . . . 21
2.1 Critérios para Contagem e Classificação de Dias Chuvosos . . . 21
2.2 Distribuições de Probabilidade Usadas na Análise de Precipitação Pluviométrica 22
2.2.1 Distribuições Probabilísticas Contínuas . . . 22
2.2.2 Distribuições Probabilísticas Discretas . . . 24
2.3 Distribuição Gama Generalizada Discreta . . . 25
2.4 Distribuição Truncada . . . 29
2.4.1 Distribuição Discreta Truncada em Zero . . . 30
2.4.1.1 Distribuição Geométrica Truncada em Zero . . . 30
2.4.1.2 Distribuição Lindley Truncada em Zero . . . 31
3 Distribuição Gama Generalizada Discreta Truncada em Zero . . . 33
4 Metodologia . . . 38
4.1 Determinação dos Dados de Estudo . . . 38
4.2 Análise dos Dados . . . 39
4.2.1 Análise Exploratória . . . 39
4.2.2 Teste Wald-Wolfowitz . . . 40
4.2.3 Teste Mann-Kendall . . . 41
4.2.4 Estimação de Máxima Verossimilhança . . . 42
4.2.5 Teste de Aderência . . . 42
4.2.6 Critérios de Informação . . . 43
4.2.6.1 Critérios de Informação de Akaike . . . 43
4.2.6.2 Critério de Informação Bayesiano . . . 43
5 Resultados e Discussões . . . 44
5.1 Análises das Séries Anuais . . . 44
5.1.1 Região Nordeste . . . 44
5.1.2 Região Norte . . . 53
5.1.3 Região Centro-Oeste . . . 59
5.1.4 Região Sudeste . . . 64

5.1.5 Região Sul 69
5.2 Análises das Séries Mensais 74
5.2.1 Estação Porto São José 75
5.2.2 Estação Porto Rico 82
6 Considerações Finais 88
7 Referências Bibliográficas 90

Capítulo 1

Introdução

A distribuição gama generalizada definida por Stacy em 1962 tem sido utilizada por muitos pesquisadores, possivelmente por ser uma distribuição flexível, devido apresentar casos especiais como a distribuição Weibull, gama, exponencial, Rayleigh, entre outras. Deste modo, ressalta-se que estudos sobre a flexibilidade da distribuição gama generalizada tem sido publicados, tais como Hasanshahi et al. (2020), Pal et al. (2019), Matheson et al. (2017). Destacando-se que Pal et al. (2019) mencionaram que a mesma permite adquirir propriedades das distribuições pertencentes aos casos particulares, e que Matheson, et al. (2017) descreveram-na também como uma distribuição robusta.

Além disso, ressalta-se que Zografos e Balakrishnan (2009) declararam que uma das razões para generalizar uma distribuição conhecida é porque a sua forma generalizada flexibiliza a distribuição de tal maneira que possa obter várias formas da função. Inclusive Achcar et al. (2017) descreveram que a distribuição gama generalizada tem desempenhado um papel importante em análises de sobrevivência, fornecendo flexibilidade e excelentes ajustes. De acordo com Pal et al. (2019), a flexibilidade dessa distribuição deriva do fato dela dispor de importantes distribuições como os casos particulares utilizadas em estudos de sobrevivência. Segundo J et al. (2019) como muitas distribuições comumente usadas para modelos paramétricos em análises de sobrevivência são casos especiais da gama generalizada, às vezes ela é utilizada para determinar qual modelo paramétrico é apropriado para um determinado conjunto de dados.

No entanto, a distribuição gama generalizada tem sido empregada em outras áreas também, como a de estudos climáticos por exemplo, Thurai e Bringi (2018) aplicaram a mesma distribuição em busca de ajuste melhores para o tamanho (diâmetro) de gotas de chuva. Assim como, Nadarajah e Gupta (2007) que usaram a distribuição para ajustar os dados do Índice de Gravidade da Seca do estado de Nebraska. Já, Papalexiou e Koutsoyiannis (2016) constaram que essa distribuição apresentou um desempenho muito bom para descrever a precipitação volumétrica diária e mensal em estações dispostas por todos os continentes.

Os pesquisadores Şen (2015), Caloiero et al. (2015), Sukla et al. (2012) e Vela et al. (2007) descreveram sobre importância de obter informações completas sobre o número de períodos de chuva e seca, em campos como, projetos de gestão de água, planejamento agrícola, estudos de enchentes e em várias atividades industriais; e além de ter o conhecimento do histórico da precipitação para o monitoramento dos impactos causados pelo excesso ou falta prolongada de precipitação, em determinada região. Assim como, a Agência Nacional das Águas - ANA (2014) declarou sobre a necessidade de medir no campo uma série de variáveis hidrológicas e meteorológicas, visando o conhecimento das características hidrológicas e possibilitando a aplicação dos modelos matemáticos na previsão de níveis e/ou vazões, a estimativa da probabilidade associada a eventos raros e a quantificação as possibilidades do aproveitamento dos recursos hídricos.

De acordo com a Organização Meteorológica Mundial - WMO (2009), os modelos probabilísticos são usados em uma ampla variedade de estudos hidrológicos, incluindo os estudos de fluxos, secas, volumes de reservatórios e quantidades de chuva. A escolha e o ajuste de uma distribuição de probabilidade que melhor represente o comportamento da variável em estudo é a parte primordial da análise. Sendo assim, inúmeras distribuições de probabilidade têm sido consideradas na literatura, entre elas, as mais utilizadas são a distribuição Pearson tipo III, Weibull, normal, log-normal, logística e Gumbel.

Deni et al. (2010), Deni e Jemain (2009) e Deni et al. (2008), em seus estudos utilizaram várias distribuições discretas de probabilidade para descreverem as características dos períodos chuvosos e secos das estações pluviométricas na Península da Malásia. Nas ultimas décadas pode-se notar um grande número de distribuições discretas publicadas e muitas das distribuições contínuas desenvolvidas no passado foram discretizadas, como por exemplo as distribuições log-logística (Para e Jan, 2016; Khorashadizadeh et al., 2013), Gama Generalizada (Chakraborty, 2015), Gumbel (Chakraborty e Chakravarty, 2014), Gama (Chakraborty e Chakravarty, 2012), Lindley (Gómez-Déniz e Calderin-Ojeda, 2011), Burr (Krishna e Pundir, 2009), Maxwell (Krishna e Pundir, 2007), Rayleigh (Roy, 2004).

Dentre a variedade de distribuições discretas propostas na literatura a distribuição Gama generalizada discreta se faz notório por dispor de dez casos particulares que inclui algumas das distribuições discretizadas citadas acima.

Diante de toda a relevância da distribuição gama generalizada discreta este trabalho buscou truncar em zero e avaliar seu desempenho no ajuste do tamanho das sequências de dias chuvosos das séries históricas anuais das estações meteorológicas das capitais do Brasil, juntamente, com as séries históricas mensais das estações meteorológicas de Porto São José e Porto Rico do rio Paraná buscando verificar se os casos particulares com 1 parâmetro da distribuição gama generalizada discreta truncada em zero apresentam bons ajustes, ou seja, analisar se as distribuições truncadas em zero gama discreta com um parâmetro e a geométrica são boas candidatas para descrever as observações da séries mensais.

1.1 Justificativa

As altas intensidades de chuvas e frequências prolongadas de dias chuvosos ou falta de chuva podem ser danoso à agricultura e catastrófico para regiões urbanas. A precipitação pluviométrica também é imprescindível para abastecer os reservatórios e as barragens de hidrelétricas, além de manter a conectividade do ecossistema aquático, como por exemplo, manter a ligação entre rios, lagos, lagoas e riachos.

É importante ressaltar que o tamanho das sequências de dias chuvosos é caracterizado como o número de dias consecutivos com precipitação superior a um limite específico, assim é notório que os dados utilizados têm truncamento em zero, logo é mais plausível trabalhar com dados truncados em zero.

Assim, devido a sua flexibilidade e por apresentar importantes distribuições como casos particulares, pode ser relevante avaliar o desempenho da distribuição gama generalizada discreta truncada em zero para descrever o tamanho das sequências de dias chuvosos.

1.2 Objetivo

Deste modo, este trabalho teve como objetivo principal avaliar o desempenho da distribuição gama generalizada discreta truncada em zero no ajuste do tamanho das sequências de dias chuvosos das séries históricas anuais das estações meteorológicas das capitais do Brasil, assim como, também, analisar as séries históricas mensais das estações meteorológicas de Porto São José e Porto Rico do rio Paraná buscando verificar se as distribuições dos casos particulares com 1 parâmetro da distribuição estudada apresentam bons ajustes aos dados. Com os seguintes objetivos específicos para as:

- séries históricas anuais:
  - analisar o tamanho das sequências do número de dias chuvosos;
  - utilizar o método da máxima verossimilhança para estimar dos parâmetros;
  - por meio das estimativas intervalares dos parâmetros identificar se há a redução da distribuição gama generalizada discreta truncada em zero em casos particulares;
  - verificar o melhor ajuste entre as distribuições dos casos particulares.
- séries históricas mensais:
  - analisar o tamanho das sequências do número de dias chuvosos;
  - avaliar o desempenho das distribuições dos casos particulares com 1 parâmetro, ou seja, analisar as distribuições gama discreta com um parâmetro truncada em zero e a geométrica truncada em zero;

- verificar o melhor ajuste entre as distribuições dos casos particulares com 1 parâmetro e Lindley discreta truncada em zero.

1.3 Organização do Trabalho

Este trabalho está organizado da seguinte forma. O primeiro capítulo com a introdução, justificativa e objetivos, como já foi apresentado. No segundo capítulo está descrito a revisão bibliográfica mencionando alguns trabalhos que descreveram a precipitação pluviométrica por meio das distribuições de probabilidade discretas e contínuas, assim como a descrição da distribuição gama generalizada discreta e apontamentos sobre a distribuição truncada em zero. No terceiro capítulo foi apresentado da função massa de probabilidade da distribuição gama generalizada discreta truncada em zero e seus casos particulares. No quarto capítulo foi registrado a metodologia das análises de dados que descrevem o tamanho das sequências de dias chuvosos. No quinto capítulo foi apresentado os resultados e discussões e por fim no ultimo capítulo as considerações finais.

Capítulo 2

Revisão Bibliográfica

2.1 Critérios para Contagem e Classificação de Dias Chuvosos

As principais características físicas dos períodos chuvosos e secos são ocasionadas pela frequência e o volume da chuva, sendo essas influenciadas pela área geográfica e época do ano (Sukla et al. 2012).

Deni et al. (2009), Chowdhury e Beecham (2013) e Rayner et al. (2016) consideraram um dia seco como aquele que apresenta uma quantidade de chuva inferior a 0,1 mm e estiagem como sendo um período de dias secos consecutivos. Já Moreira et al., (2010) consideram dias secos com limiar igual ou inferior a 5 mm/dia. Já Sukla et al. (2012) definiram um dia seco como um dia com menos de 2,5 mm de chuva, caso contrário era considerado um dia chuvoso ($\geq$ 2,5 mm)

De acordo com Mandapaka et al. (2016) o período chuvoso é definido como o número de dias consecutivos com precipitação maior ou igual a um limite específico e utilizaram os limiares 0,1 e 1 mm/dia em seus estudo. Ao passo que Silva et al., (2011), consideram o dia chuvoso aquele com precipitação acima de 0,1 mm. Oliveira (2016) adotou como um dia chuvoso os dias com precipitação superior ou igual 5 mm/dia.

Zolina et al. (2013), ressaltaram que o uso de séries temporais incluindo precipitação leve desafia a robustez dos resultados. Além dos altos erros observacionais aleatórios para a precipitação leve, as práticas observacionais de precipitação leve também diferem entre os países europeus, o que pode resultar em vieses regionais artificiais em conjunto de dados multinacionais.

Para o cálculo da frequência das precipitações, Santos e Ferreira (2016) contaram o número total de dias em que ocorreram chuvas, de acordo com quatro classes pré-estabelecidas: $<$ 10 mm; 10.1 mm a 20 mm; 20.1 mm a 29.9 mm e $>$ 30.0 mm. Assim, as chuvas foram classificadas conforme exposto na Tabela 1.

Tabela 1 – Tipo de precipitação de acordo com a intensidade.

Classe de Precipitação diária	Tipo de Precipitação
0 – 10.0 mm	Chuva Fraca
10.1 – 20.0 mm	Chuva Moderada
20.1 – 30.0 mm	Chuva Moderada a Forte
$>$ 30.0 mm	Chuva Forte

Fonte: Santos e Ferreira (2016)

Já Gouvea et al. (2018) decidiram pela classificação da precipitação diária dividida em: chuvisco (0,1 a 2,5 mm), chuva fraca (2,5 a 10,0 mm), chuva moderada (10,0 a 15,0 mm), chuva relativamente forte (15 a 25mm), chuva forte (25,0 a 50,0 mm) e chuva extrema (acima de 50 mm), para analisar a frequência de precipitação na bacia no Rio Itajaí - SC.

2.2 Distribuições de Probabilidade Usadas na Análise de Precipitação Pluviométrica

A fonte natural mais importante de água para os seres vivos é chuva que contribui para abastecimento dos lagos e rios. Várias distribuições de probabilidade têm sido aplicadas para descrever a precipitação observada em diferentes períodos, uma vez que a precipitação varia com o tempo e a localização (Amin et al. 2016).

Na literatura encontram-se muitos estudos sobre a modelagem da precipitação pluviométrica e alguns deles são citados a seguir.

2.2.1 Distribuições Probabilísticas Contínuas

Inúmeros são as distribuições de probabilidades utilizadas na análise de precipitação e alguns estudos foram citados como os de Husak et al. (2007) que utilizaram a distribuição gama para descrever a pluviosidade mensal na África de 1961 a 1996, uma região bastante heterogênea do ponto de vista atmosférico, e concluíram que essa distribuição apresenta um bom ajuste aos variados conjuntos de dados observados de chuva.

Em um estudo na África do Sul, Mzezewa et al. (2010) analisaram as séries anuais e mensais dos dados de precipitação da estação meteorológica na Província de Limpopo empregando as distribuições normal, log-normal, gama e Weibull. Os pesquisadores constataram que a precipitação anual foi melhor descrita pela distribuição normal e a estação chuvosa foi melhor descrita pela distribuição log-normal nos meses de outubro, novembro e dezembro, já os meses de janeiro e fevereiro pela distribuição Weibull e o mês de março pela distribuição Gamma.

Ramos e Moala (2014) estudaram as séries históricas da estacão meteorológica de Presidente Prudente analisando os dados de precipitação total mensal de janeiro de 1943 a dezembro de 2003 e verificaram que distribuição exponencial geométrica estendida apresentou um melhor ajuste em relação as distribuições Gama, Weibull e log-normal.

Amin et al. (2016) com o objetivo de estudar os valores máximos anuais de precipitação em 6 estações meteorológica do norte do Paquistão de 1961 a 2010. Os pesquisadores empregaram as distribuições normal, log-normal, log-Pearson tipo III e Gumbel para obter o melhor ajuste. A distribuição normal forneceu o melhor ajuste para estação de chuvas de Mardan, enquanto que a distribuição log-Pearson tipo III apresentou o melhor ajuste para às outras estações de chuvas analisadas (Besham Qilla, Daggar, Kalam, Oghi e Puran).

Kumar et al. (2017) estudaram as séries anuais dos dados de precipitação de 13 distritos do estado de Uttarakhand - Índia nos períodos de 1991 a 2002, com as distribuições: qui-quadrado, qui-quadrado (2 parâmetros (2P)), exponencial, exponencial (2P), gama, gama (3P), valores extremos generalizados, Pearson tipo III, Weibull, Weibull (3P). E concluíram que as distribuições Weibull, Weibull (3P), Qui-quadrado (2P), gama e log-Pearson (3P) foram as mais adequadas para 31%, 15%, 24%, 15% e 15% das estações estudadas, respectivamente.

Cremoneze et al. (2017) com objetivo de selecionar uma distribuição de probabilidade que descreva o comportamento das precipitações pluviométricas máximas mensais de 26 estações meteorológicas localizadas nos estados do Paraná e Rio Grande do Sul no período de 1961 a 2015, foi constatado que a distribuição Gumbel obteve o melhor ajuste entre as distribuições Gumbel transmutada, Gumbel Marshall-Olkin, Gumbel exponenciada, Gumbel generalizada e Exponencial-Gama com três parâmetros.

Reis (2017) analisou a precipitação acumulada mensal das séries de 33 estações meteorológicas da região Sul do Brasil nos períodos de 1961 a 2016 e verificou que a distribuição Weibull foi a que apresentou melhor performance no ajuste dos dados em relação as distribuições Odd Weibull, Marshall-Olkin Weibull, Weibull exponenciada e Weibull transmutada e gama.

Ye et al. (2018) analisaram os recordes pluviométricos dos dias chuvosos de 237 estações dos Estados Unidos no período de 1948 a 2003 em 49 estados. E ressaltaram que a distribuição Gama com dois parâmetros mostrou-se uma distribuição adequada aos dados, porém às distribuições de Pearson Tipo III e Kappa apresentaram os melhores ajustes em relação a distribuição gama e as outras oito distribuições analisadas (Weibull, Pearson Tipo III, logística generalizada, valor extremo generalizado Tipo III, Pareto generalizada com 2 e 3 parâmetros, log-normal com 2 e 3 parâmetros).

Mazucheli e Emanuelli (2019) avaliaram as séries históricas mensais acumuladas de 33 estações meteorológicas da região Sul do Brasil de 1970 a 2014 e constaram que a distribuição Nakagami obteve o melhor desempenho em relação as distribuições Weibull, gama, log-normal, log-logística e inversa-gaussiana.

2.2.2 Distribuições Probabilísticas Discretas

Muitas funções massa de probabilidade foram publicadas e algumas estão sendo empregadas no estudo do comportamento dos períodos chuvosos e secos.

Deni et al. (2008) examinaram o ajuste das distribuições geométrica, geométrica composta, séries logarítmicas, logarítmica modificada, Polia, log-geométrica e binomial negativa truncada para os períodos de chuva e de seca da precipitação diária em 10 estações pluviométricas na Península da Malásia observados entre 1971 e 2005. E concluíram que distribuição geométrica composta e a distribuição binomial negativa truncada foram as distribuições mais adequadas para descrever as características dos períodos chuvosos e secos, respectivamente. Este estudo também identificou que a distribuição de probabilidade mais apropriada para cada estação pluviométrica na Península da Malásia, variou de acordo com a localização das estações.

Deni e Jemain (2009) propuseram cinco diferentes misturas com as distribuições séries de logarítmicas, geométrica, Poisson e Poisson truncada para descrever a ocorrência de eventos diários dos períodos chuvosos e secos das 14 estações pluviométricas selecionadas na Peninsular Malásia períodos de 1975 a 2004.

Deni et al. (2010) descreveram que das 13 distribuições avaliadas, a mistura das distribuições logarítmica e Poisson truncada foi a que apresentou o melhor desempenho nos comportamento dos tamanhos das sequências dos períodos secos e chuvosos para maioria das 38 estações pluviométricas da Península da Malásia no período de 1975 a 2004.

Mathugama e Peiris (2011) utilizaram os modelos de Markov e várias distribuições de probabilidade tais como binomial negativa, binomial negativa truncado, Polia-Eggemberg. Estes autores concluíram que a distribuição binomial negativa apresentou melhor desempenho do que os modelos de Markov para explicar frequências de seca na base sazonal de Sri Lanka.

Sukla et al. (2012), com o objetivo de identificar a melhor distribuição que ajusta o tamanho das sequências de dias chuvosos e secos compararam o modelo da Cadeia de Markov e as distribuições binomial negativa truncada e a Eggemberger-Polya em quatro estações meteorológicas da região do delta de Mahanadi na Índia (Bhubaneswar, Cuttack, Paradip e Puri), concluíram pelo modelo mais simples utilizando o princípio da parcimônia, isto é, Cadeia de Markov baseado na distribuição geométrica.

Zolina et al. (2013) analisaram as mudanças na duração dos períodos secos e chuvosos na Europa durante o período de 1950 a 2009, com as distribuições geométrica truncada e geométrica truncada fracionada.

Oliveira (2016) estudou o tamanho das sequências de dias chuvosos de 33 estações da região Sul do Brasil num período de 1995 a 2014 e concluiu que a distribuição Lindley discreta truncada em zero obteve o melhor ajuste em relação às outras 11 distribuições avaliadas, seguido da distribuição binomial negativa truncada em zero.

2.3 Distribuição Gama Generalizada Discreta

Chakraborty (2015) discretizou a distribuição gama generalizada (Stacy, 1962) via a função de sobrevivência, os pioneiros desse método foram Toshio Nakagawa e Shunji Osaki em 1975, vale ressaltar que é um dos métodos mais utilizados na literatura.

Como descrito por Chakraborty e Chakravarty (2012), dada qualquer distribuição contínua por ser discretizada, se Y é uma variável aleatória contínua tem função de sobrevivência definida por $S_Y(y) = P(Y \geq y)$, então a variável aleatória $X = \lfloor Y \rfloor$, em que $\lfloor Y \rfloor$ indica a parte inteira menor ou igual a Y, tem a função definida por:

$$\begin{aligned} P(X = x) &= P(x \leq Y < x+1) = P(Y \geq x) - P(Y \geq x+1) \\ &= S_Y(x) - S_Y(x+1), \qquad x = 0, 1, 2, \ldots \end{aligned} \tag{2.1}$$

Assim, tendo que a função densidade de probabilidade da distribuição gama generalizada é representada por:

$$f(y \mid k, \theta, c) = \frac{c}{\theta^{ck}\Gamma(k)} y^{(ck-1)} e^{-(y/\theta)^c} , \tag{2.2}$$

sendo $y \geq 0$, $\theta > 0$ o parâmetro escala, $k > 0$ e $c > 0$ são os parâmetros de forma.

A função de sobrevivência é dada por

$$S_Y(y \mid k, \theta, c) = \frac{1}{\Gamma(k)} \left[\Gamma\left(k, \left(\frac{y}{\theta}\right)^c\right) \right] \tag{2.3}$$

sendo $\Gamma(k\,, a) = \int_a^\infty v^{(k-1)} e^{-(v/\theta)} dv$ é denominada de função gama incompleta superior.

Substituindo (2.3) em (2.1) tem-se que a versão da distribuição discreta, isto é, uma variável aleatória discreta X tem distribuição gama generalizada discreta se sua função massa de probabilidade é dada por:

$$\begin{aligned} P(X = x \mid k, \theta, c) &= \frac{1}{\Gamma(k)} \left[\Gamma\left(k, \left(\frac{x}{\theta}\right)^c\right) - \Gamma\left(k, \left(\frac{x+1}{\theta}\right)^c\right) \right] \\ &= \frac{1}{\Gamma(k)} \left[\Gamma\left(k\,, \left(\frac{x}{\theta}\right)^c, \left(\frac{x+1}{\theta}\right)^c\right) \right] \end{aligned} \tag{2.4}$$

em que:

$x = 0, 1, 2, \ldots$; $\theta > 0$; $k > 0$ e $c > 0$;

$\Gamma\left(k\,, \left(\frac{x}{\theta}\right)^c, \left(\frac{x+1}{\theta}\right)^c\right) = \frac{c}{\theta^{ck}} \int_x^{x+1} u^{(ck-1)} e^{-(u/\theta)^c} du$;

$\Gamma\left(k\,, \left(\frac{x}{\theta}\right)^c\right) = \frac{c}{\theta^{ck}} \int_x^{\infty} u^{(ck-1)} e^{-(u/\theta)^c} du$;

$\Gamma\left(k\,, \left(\frac{x+1}{\theta}\right)^c\right) = \frac{c}{\theta^{ck}} \int_{(x+1)}^{\infty} u^{(ck-1)} e^{-(u/\theta)^c} du$.

A distribuição gama generalizada discreta é denotada por $GaGD(k, \theta, c)$ e sua função de distribuição acumulada é definida por:

$$F(x \mid k, \theta, c) = 1 - \left[\frac{1}{\Gamma(k)}\Gamma\left(k, \left(\frac{x}{\theta}\right)^c\right)\right], \tag{2.5}$$

uma vez que, Chakraborty (2015), declarou que a distribuição discreta detém a mesma forma funcional da função de sobrevivência do caso contínuo, isto é, $S_X(x) = S_Y(y)$.

Para uma variável aleatória tem-se o momento de ordem r e pode ser escrito na forma:

$$E(X^c) = \sum_{i=1}^{r}(-1)^{i+1}\binom{r-1}{i}g^{r-i}\left(k\,,\left(\frac{x}{\theta}\right)^c\right) + g^{r-1}\left(k\,,\left(\frac{x}{\theta}\right)^c\right) - E(X^{r-1}) \tag{2.6}$$

em que:

$$g^r\left(k\,,\left(\frac{x}{\theta}\right)^c\right) = \frac{1}{\Gamma(k)}\sum_{x=1}^{\infty} x^r\,\Gamma\left(k\,,\left(\frac{x}{\theta}\right)^c\right). \tag{2.7}$$

Assim, os dois primeiros momentos é dada por:

$$E(X) = \frac{1}{\Gamma(k)}\sum_{x=1}^{\infty}\Gamma\left(k\,,\left(\frac{x}{\theta}\right)^c\right) \tag{2.8}$$

$$E(X^2) = \frac{2}{\Gamma(k)}\sum_{x=1}^{\infty} x\,\Gamma\left(k\,,\left(\frac{x}{\theta}\right)^c\right) - E(X) \tag{2.9}$$

A Figura 1 mostra o comportamento da função massa de probabilidade da distribuição $GaGD(k, \theta, c)$. Ressalta-se, ao observar as Figuras 1a e 1b, é notório que essa distribuição tem curva decrescente, com a moda centrada em zero . Comparando as Figuras 1c e 1d percebe-se que o parâmetro k influencia no achatamento das curvas.

A principal motivação para se trabalhar com a distribuição $GaGD(k, \theta, c)$ é devido ter casos particulares de importantes distribuições, tais como:

i. Quando $c = 1$, tem-se a distribuição gama discreta (Chakraborty e Chakravarty, 2012), com a função massa de probabilidade é representada por:

$$P(X = x \mid k, \theta) = \frac{1}{\Gamma(k)}\left[\Gamma\left(k, \left(\frac{x}{\theta}\right)\right) - \Gamma\left(k, \left(\frac{x+1}{\theta}\right)\right)\right]. \tag{2.10}$$

Abouammoha e Alhazzani (2015) declararam que a distribuição $GaD(k, \theta)$ podem lidar tanto com dados com sub-dispersão ou com super-dispersão.

ii. Para $c = 1$ e $\theta = 1$, ela se reduz para distribuição gama discreta de um parâmetro (Chakraborty e Chakravarty, 2012), isto é,

$$P(X = x \mid k) = \frac{1}{\Gamma(k)}\left[\Gamma\big(k, x\big) - \Gamma\big(k, (x+1)\big)\right]. \tag{2.11}$$

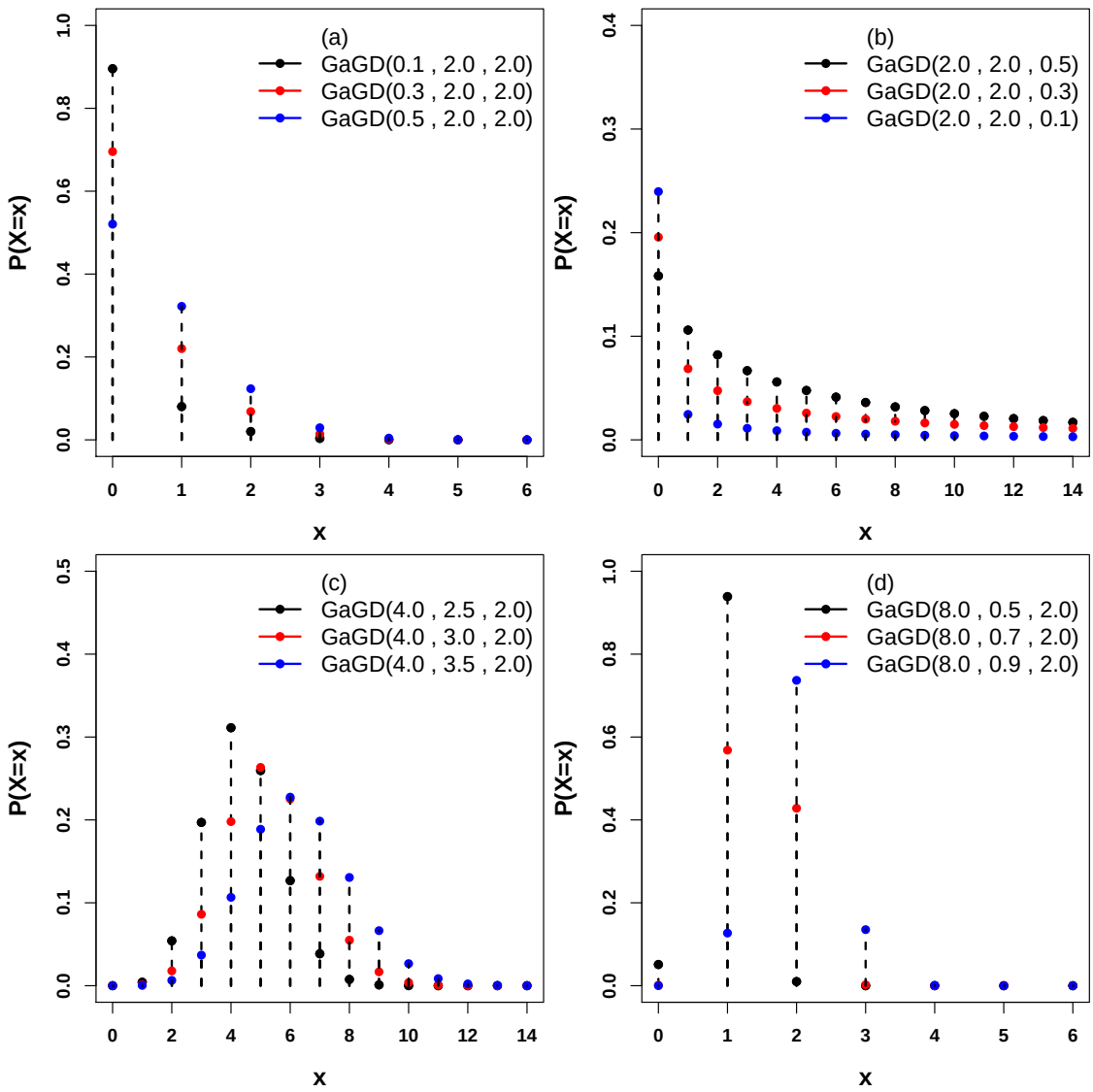

Figura 1 – Comportamento da função massa de probabilidade de $GaGD(k, \theta, c)$.

iii. Quando $k = 1$, a distribuição reduz a distribuição Weibull discreta (Nakagawa and Osaki, 1975), então a função de massa de probabilidade é dada por:

$$P(X = x \mid c\ ,\ q) = q^{x^c} - q^{(x+1)^c}\ ; \quad x = 0, 1, 2, \cdots\ ;\ c > 0\ ;\ 0 < q < 1 \tag{2.12}$$

em que $q = e^{-(1/\theta)^c}$.

iv. Se os parâmetros forem $c = 1$ e $k = 1$, essa distribuição se torna a distribuição geométrica, isto é,

$$P(X = x \mid q) = q^{x} - q^{(x+1)}\ ; \quad x = 0, 1, 2, \cdots\ ;\ 0 < q < 1\,, \tag{2.13}$$

em que $q = e^{-(1/\theta)}$.

v. Com $c = 2$ e $k = 1$ obtêm-se a distribuição Rayleigh discreta (Roy, 2004) com função massa de probabilidade representada por:

$$P(X = x \mid q) = q^{x^2} - q^{(x+1)^2}\ ; \quad x = 0, 1, 2, \cdots\ ;\ 0 < q < 1\,, \tag{2.14}$$

em que $q = e^{-(1/\theta)^2}$.

vi. Para $c = 2$ e $k \longrightarrow \dfrac{k}{2}$, a distribuição se torna a distribuição Rayleigh generalizada discreta (Alamatsaz et al., 2016), isto é,

$$P(X = x \mid k\ ,\ \theta) = \frac{1}{\Gamma\left(\frac{k}{2}\right)}\left[\Gamma\left(\frac{k}{2}, \left(\frac{x}{\theta}\right)^2, \left(\frac{x+1}{\theta}\right)^2\right)\right] \tag{2.15}$$

$$x = 0, 1, 2, \cdots\ ;\ k > 0\ ;\ \theta > 0\ .$$

vii. Tomando $c = 2$, $GaGD(k, \theta, c)$, reduz a distribuição hidrograph discreta (Chakraborty, 2015), ou seja,

$$P(X = x \mid k\ ,\ \theta) = \frac{2}{\theta^{2k}\Gamma(k)}\int_x^{x+1} u^{(2k-1)}e^{-(u/\theta)^2}du \tag{2.16}$$

$$x = 0, 1, 2, \cdots\ ;\ k > 0\ ;\ \theta > 0\ .$$

viii. Para $c = 2$, $k = \dfrac{3}{2}$ e $\theta \longrightarrow \sqrt{\theta}$, a distribuição se reduz a Maxwell-Boltzmann discreta (Krishna e Pundir, 2007), com função massa de probabilidade formulada por:

$$P(X = x \mid \theta) = \frac{2}{\sqrt{\pi}}\left[\Gamma\left(\frac{3}{2}, \left(\frac{x}{\theta}\right)^2, \left(\frac{x+1}{\theta}\right)^2\right)\right] \tag{2.17}$$

$$x = 0, 1, 2, \ldots\ ;\ k > 0\ ;\ \theta > 0\ .$$

ix. Com $c = 2$ e $k = \dfrac{1}{2}$, essa distribuição se torna a distribuição half-normal discreta estudada por Kemp (2008), isto é, a função massa de probabilidade é dada por:

$$\begin{aligned} P(X = x \mid \theta) &= \frac{1}{\Gamma\left(\frac{1}{2}\right)}\left[\Gamma\left(\frac{1}{2}, \left(\frac{x}{\theta}\right)^2\right) - \Gamma\left(\frac{1}{2}, \left(\frac{x+1}{\theta}\right)^2\right)\right] \\ &= 2\left[\Phi\left(b\sqrt{2}\right) - \Phi\left(a\sqrt{2}\right)\right], \end{aligned} \tag{2.18}$$

em que $a = \dfrac{x}{\theta}$ e $b = \dfrac{(x+1)}{\theta}$; $x = 0, 1, 2, \ldots\ ; b > a > 0\ ;\ \theta > 0\ .$

x. Para $k \longrightarrow \infty$, tomando $\mu = log\,\theta + \left(\dfrac{1}{c}\right)log\,k$ e $\sigma = \left(\dfrac{1}{c}\right)\sqrt{k}$, a distribuição reduz a distribuição log-normal discreta (Bi et al., 2001), isto é,

$$P(X = x \mid \mu\ ,\ \sigma) = \Phi\left[\frac{log(x+1) - \mu}{\sigma}\right] - \Phi\left[\frac{log(x) - \mu}{\sigma}\right] \tag{2.19}$$

$$x = 0, 1, 2, \ldots\ ;\ \mu > 0\ ;\ \sigma > 0\ .$$

Considerando uma amostra aleatória $x_1, x_2, \ldots, x_n$ e a distribuição $GaGD(k, \theta, c)$ como candidatas as estimação, então a função de verossimilhança é dada por

$$L(k, \theta, c \mid \mathbf{x}) = \left[\frac{1}{\Gamma(k)}\right]^n \prod_{i=1}^{n} \left[\Gamma\left(k, \left(\frac{x}{\theta}\right)^c, \left(\frac{x+1}{\theta}\right)^c\right)\right] \tag{2.20}$$

aplicando o logaritmo natural em (2.20) tem-se a função de log verossimilhança

$$l(k, \theta, c \mid \mathbf{x}) = -n \log \Gamma(k) + \sum_{i=1}^{n} \log \left[\Gamma\left(k, \left(\frac{x}{\theta}\right)^c, \left(\frac{x+1}{\theta}\right)^c\right)\right] \tag{2.21}$$

Como as equações de verossimilhança não têm solução de forma fechada, há necessidade de um método numérico para encontrar a solução deste sistema não linear.

2.4 Distribuição Truncada

Muitas vezes em diversas situações, existem análises que a exclusão da não ocorrência do fenômeno probabilístico pode originar melhores resultados, como por exemplo, no estudo sobre número de dias que o paciente ficou internado numa clínica de reabilitação, ou o número de itens na cesta dos clientes na fila do caixa rápido do supermercado, ou o tamanho das sequências de dias chuvosos no ano num ambiente tropical. Esses exemplos têm algo em comum, pois os mesmos não têm incluso o valor zero, ou seja, o paciente ficou internado pelo menos um dia na clínica, o cliente ficará na fila do pagamento se tiver no mínimo um item para comprar e ao menos terá um dia com chuvas sucessivas no ano.

Assim, ao retirar da distribuição em estudo à ocorrência de zero ou algum limiar máximo ou mínimo que não pertence ao objetivo do estudo tem-se a distribuição truncada, ou seja, o truncamento é uma característica da distribuição de probabilidade da qual as observações provêm da não ocorrência de eventos que não sejam adequadamente modelados pela distribuição, então se removê-la tem-se o modelo truncado (Greene, 2012).

Segundo Boswell et al. (1979), seja X uma variável aleatória discreta (contínua) e se a distribuição é truncada de modo que somente os valores em $w = \{a, a+1, \ldots, b\}$ são observados, então a função densidade de probabilidade da variável aleatória truncada é dada por

$$f(x) = \frac{f(x)}{\int_a^b f(x)\, dx} \tag{2.22}$$

e a função massa de probabilidade truncada é representada por

$$P(X = x) = \frac{P(X = x)}{\sum_{x=a}^{b} P(X = x)} \tag{2.23}$$

quando $a = 0$ a distribuição é chamada de truncada a direita; quando $b \to \infty$ é denominado de truncada a esquerda e quando $a = 1$ e $b \to \infty$ a distribuição é a truncada em zero.

2.4.1 Distribuição Discreta Truncada em Zero

A forma mais comum de truncamento é a omissão do valor zero. A solução de modificar a distribuição excluindo as observações com zero é chamado de distribuição truncada em zero (Zuur et al., 2009). As distribuições truncadas em zero são extensões importantes das distribuições discretas (Yee, 2015).

Assim, segundo Boswell et al. (1979), quando $a = 1$, $b \to \infty$ e X uma variável aleatória discreta tem-se que a função massa de probabilidade da distribuição truncada em zero é dada por

$$P(X = x) = \frac{P(X = x)}{\sum_1^\infty P(X = 1)} = \frac{P(X = x)}{P(X \geq 1)} = \frac{P(X = x)}{1 - P(X = 0)} \tag{2.24}$$

em que $P(X = x)$ é a função massa de probabilidade da distribuição em estudo e $x = 1, 2, ...$.

Contar dados sem a categoria zero geralmente ocorre em vários campos, como da saúde, humanas, ciências e engenharias. Uma classe de distribuições discretas truncadas em zero, como as distribuições: Poisson, geométrica e binomial negativa são propostas na literatura para modelar os dados de contagem (Tian et al., 2019).

Versões truncadas de distribuição de probabilidade discretas têm sido empregadas em estudos de precipitação pluviométrica, já mencionados os pesquisadores Deni et al. (2008) e Mathugama e Peiris (2011) que utilizaram a distribuição binomial negativa truncada; Zolina et al. (2013) com as distribuições geométrica truncada e geométrica truncada fracionada. Oliveira (2016) empregou as distribuições truncadas em zero: Lindley potência discreta, Lindley discreta, binomial negativa e Poisson.

2.4.1.1 Distribuição Geométrica Truncada em Zero

A distribuição geométrica é uma das distribuições importantes empregada na área de confiabilidade / análise sobrevivência e teoria das filas (Devi et al., 2017). Já foi chamada de distribuição Furry (Johnson et al., 2005).

Um dos primeiros relatos publicados sobre a distribuição geométrica truncada foi com Thomasson e Kapadia em 1968, nesse estudo eles recomendaram usar amostras de tamanho relativamente grande para reduzir a variação do estimador. Zolina et al. (2013), empregaram a distribuição geométrica truncada nas análises de períodos secos e chuvosos.

Segundo Klugman et al. (2012), a distribuição geométrica truncada em zero, denotada por $GeoTZ(q)$, tem função massa de probabilidade dada por

$$P(X = x \,|\, q) = (1 - q)q^{x-1} = pq^{x-1} \tag{2.25}$$

para $x = 1, 2, ...$ e $0 < q < 1$.

Ressalta-se que para alguns pesquisadores, como por exemplo Deni et al. 2008, uma variável aleatória com distribuição geométrica definida em $x = 1, 2, ..., n$ tem função massa de probabilidade dada como a (2.25).

Na representação gráfica da função massa de probabilidade da distribuição geométrica truncada em zero apresentada na Figura 2 pode-se notar a moda 1 e curva decrescente para diferentes parâmetros avaliados. Na Figura 2a pode-se observar que exatamente no ponto 8 a probabilidade é próxima de zero para o parâmetro q inferior ou igual a 0.5, isto é, $P(X = 8\,|\,0.5) \cong 0$, $P(X = 8\,|\,0.3) \cong 0$, $P(X = 8\,|\,0.1) \cong 0$.

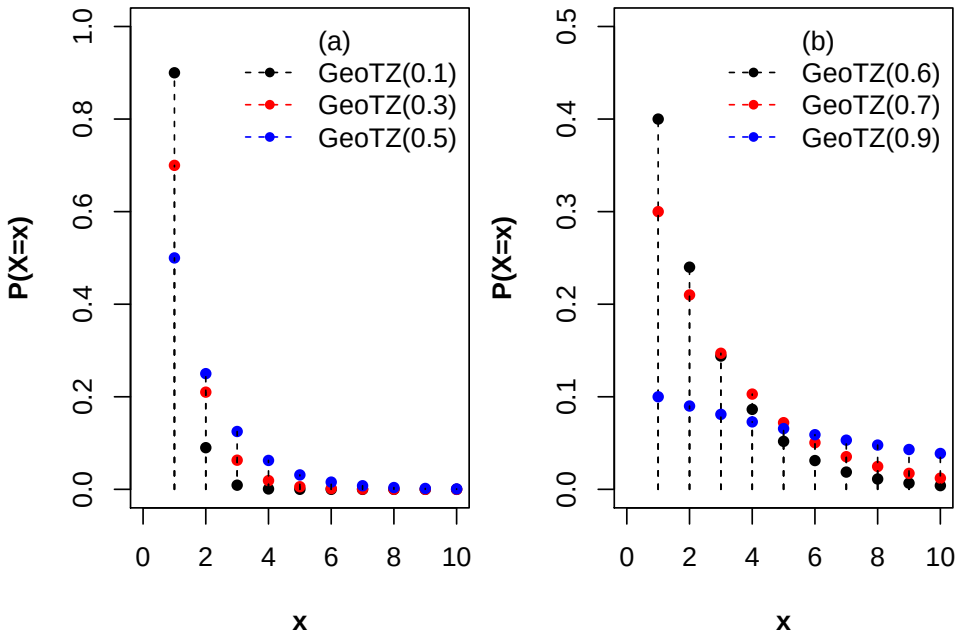

Figura 2 – Comportamento da função massa de probabilidade da distribuição $GeoTZ(q)$.

O valor esperado dessa distribuição truncada em zero é dado por

$$E(x) = \frac{1}{1-q} \tag{2.26}$$

e a variância é obtida por

$$Var(x) = \frac{q}{(1-q)^2} \tag{2.27}$$

2.4.1.2 Distribuição Lindley Truncada em Zero

Os pesquisadores Emilio Gómez-Déniz e Enrique Calderín-Ojeda em 2011 usando o método de discretização proposto por Nakagawa e Osaki (1975) apresentaram uma versão discreta da distribuição Lindley. Os mesmos afirmaram que uma das vantagens da sua utilização ocorre devido ao fato da distribuição ter superdispersão (variância maior que a média) sendo, portanto, mais flexível que a distribuição de Poisson para modelar dados de contagens.

Oliveira (2016) propôs truncar as distribuições Lindley discreta, Lindley Potência Discreta, Lindley Potência Discreta Deslocada da Origem e Lindley Ponderada Discreta e utilizou tais distribuições para verificar adequabilidade e o melhor desempenho entre elas para descrever o tamanho das sequências de dias chuvosos de 33 estações da região Sul do Brasil, concluindo que distribuição Lindley discreta truncada em zero obteve o melhor ajuste em termos de parcimônia. Assim, truncando em zero a função proposta por Bakouch et al. (2014) tem-se que a função massa de probabilidade Lindley discreta truncada em zero é dada por:

$$P(X = x \mid \beta) = \frac{e^{-\beta x}[\beta(1 - 2e^{-\beta}) + (1 - e^{-\beta})(1 + \beta x)]}{e^{-\beta}(1 + 2\beta)}, \tag{2.28}$$

em que $x = 1, 2, 3, \ldots$; $\beta > 0$. A distribuição truncada é denotada por $LiDTZ(\beta)$ tem apenas uma moda.

O comportamento da função massa de probabilidade da distribuição $LiDTZ(\beta)$ para algumas configurações estão apresentados na Figura 3 e pode-se notar que para os valores do parâmetro $\beta = 3.0$, $\beta = 2.0$, $\beta = 1.0$, $\beta = 0.9$ e $\beta = 0.6$ as modas foram iguais 1 com curvas decrescentes, enquanto que para $LiDTZ(0.3)$ a curva foi assimétrica com moda igual a 2.

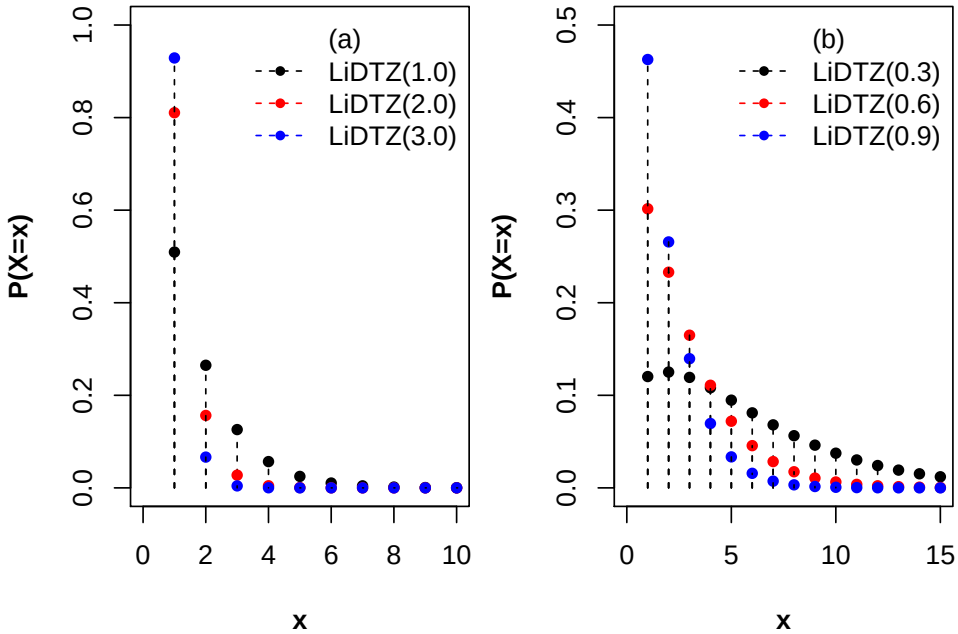

Figura 3 – Comportamento da função massa de probabilidade da distribuição $LiDTZ(\beta)$.

A média é descrita por:

$$E(X) = \frac{2\beta - \beta e^{-\beta} + 1 - e^{-\beta}}{(1 + 2\beta)(1 - e^{-\beta})^2} \tag{2.29}$$

e a variância por

$$Var(X) = \frac{(\beta e^{-\beta} - 4\beta - 1)e^{-\beta}}{(1 + 2\beta)(1 - e^{-\beta})^3} \frac{E(X)}{e^{-\beta} - 1} \tag{2.30}$$

Capítulo 3

Distribuição Gama Generalizada Discreta Truncada em Zero

Na teoria das probabilidades, as distribuições com truncamento em zero são distribuições discretas cujo suporte é o conjunto de inteiros positivos. A distribuição truncada em zero é a escolha apropriada quando os dados a serem modelados são originários de mecanismo que gera dados que excluem estruturalmente a contagem zero (Shanker e Shukla, 2017).

Assim, para estudar o tamanho das sequências de dias chuvosos com a distribuição gama generalizada discreta propôs-se truncar em zero essa distribuição, pois na contagem da sequência de dias chuvosos o zero não faz parte da amostragem. Haja vista que o fato de não chover (o zero ou limiar inferior ao determinado) é a ocorrência que faz parte da contagem das sequências de dias secos.

Diante disso, tomando a função massa de probabilidade (2.4) e truncando em zero a distribuição $GaGD(k, \theta, c)$ conforme a função (2.24) tem-se que

$$P(X = x \mid k\ , \theta\ ,\ c) = \frac{\Gamma\left(k, \left(\frac{x}{\theta}\right)^c\right) - \Gamma\left(k, \left(\frac{x+1}{\theta}\right)^c\right)}{\Gamma\left(k, \left(\frac{1}{\theta}\right)^c\right)} \tag{3.1}$$

$$x = 1, 2, 3, \cdots\ ;\ k > 0\ ;\ \theta > 0 \quad \text{e} \quad c > 0.$$

A distribuição gama generalizada discreta truncada em zero é denotada por $GaGDTZ(k, \theta, c)$.

Observando a Figura 4, para as configurações apresentadas, pode-se notar que o comportamento da função massa de probabilidade da distribuição $GaGDTZ(k, \theta, c)$ é unimodal. Salientando que para os valores dos parâmetros estabelecidos nas Figuras 4a, 4d e 4g a distribuição da curva é decrescente com a moda centrada em um, enquanto que nas outras

observou-se uma curva assimétrica positiva e apresentando o mesmo contorno da distribuição $GaGD(k, \theta, c)$, porém iniciando no valor 1.

Ressalta-se também a flexibilidade da distribuição, por exemplo, para os parâmetros estabelecidos na Figura 4a, exatamente, no ponto 5 a probabilidade está mais próxima de zero, isto é, $P(X = 5 \mid k = 1\,, \theta = 2\,, c = 2) \cong 0$, enquanto que no ponto 5 da Figura 4d a probabilidade está acima de 1%, ou seja, $P(X = 5 \mid k = 1\,, \theta = 1\,, c = 1) > 1\%$, já na Figura 4g tem-se no ponto 5 uma probabilidade acima de 4%, isto é, $P(X = 5 \mid k = 1\,, \theta = 0.5\,, c = 0.5) > 4\%$, tais ocorrências são importantes na descrição do tamanho de sequências de dias chuvosos devido a diversidade climática e consequentemente a amplitude do tamanho dessas sequências diferem entre as estações meteorológicas.

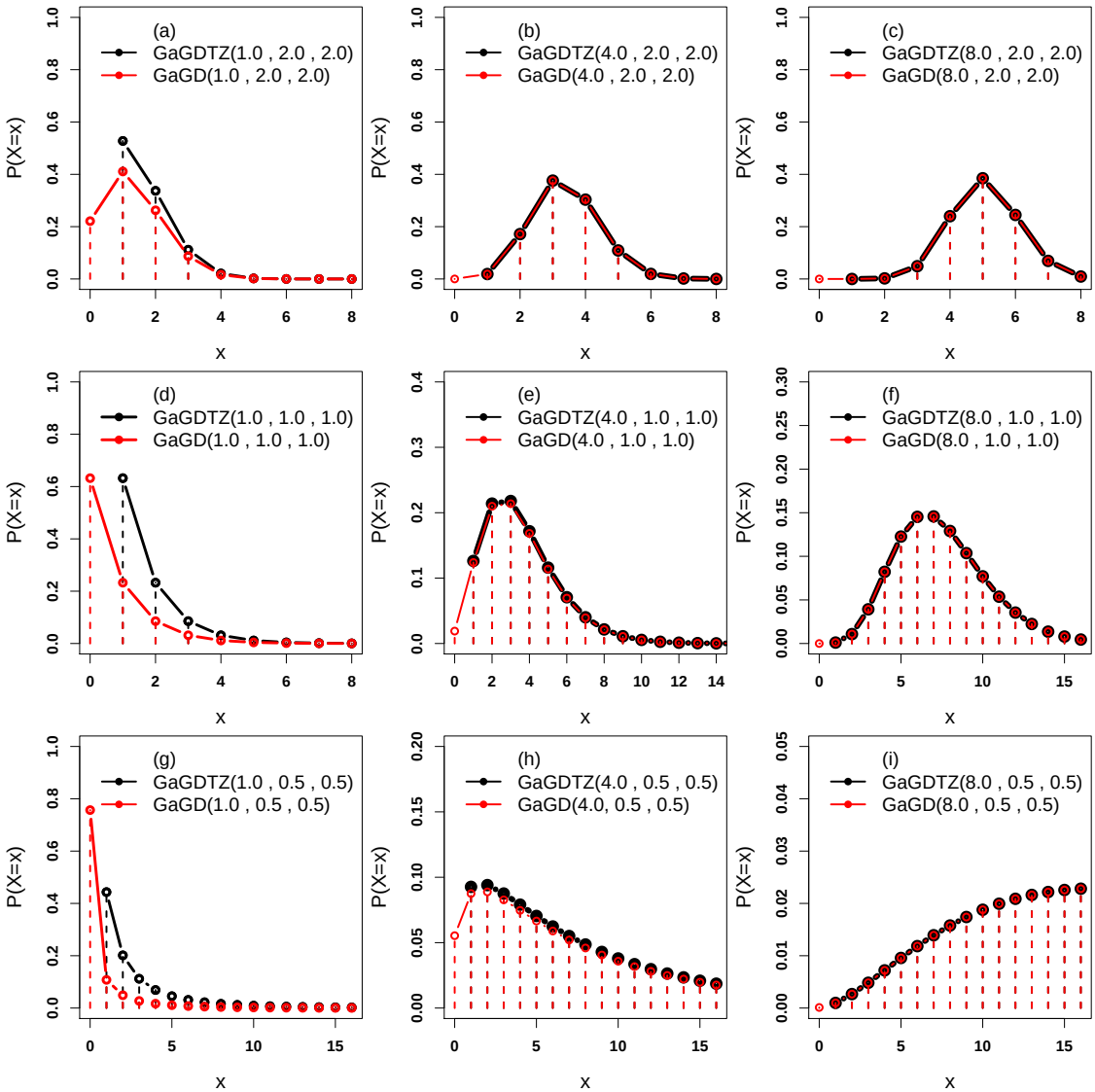

Figura 4 – Comportamento da funções massa de probabilidade $GaGDTZ(k, \theta, c)$ e $GaGD(k, \theta, c)$.

Os casos particulares da distribuição $GaGDTZ(k, \theta, c)$ são os mesmos da distribuição sem truncamento como seguem abaixo:

i. Com $c = 1$, essa distribuição se reduz a distribuição gama discreta truncada em zero, isto é

$$P(X = x \mid k\,, \theta) = \frac{\Gamma\left(k\,, \left(\frac{x}{\theta}\right)\right) - \Gamma\left(k\,, \left(\frac{x+1}{\theta}\right)\right)}{\Gamma\left(k\,, \left(\frac{1}{\theta}\right)\right)}, \tag{3.2}$$

$x = 1, 2, 3, \ldots\,;\ k > 0\,; \theta > 0$. Tal distribuição será denotada por $GaDTZ(k, \theta)$.

ii. Para $c = 1$ e $\theta = 1$, ela se reduz a distribuição gama discreta truncada em zero com um parâmetro, representada por:

$$P(X = x \mid k) = \frac{\Gamma\big(k\,, x\big) - \Gamma\big(k, x+1\big)}{\Gamma(k)}\,, \tag{3.3}$$

$x = 1, 2, 3, \ldots\,;\ k > 0\,;$ com a distribuição denotada por $GaDTZ(k)$.

iii. Com $k = 1$, a distribuição reduz a distribuição Weibull discreta truncada em zero, então tem-se que:

$$P(X = x \mid c\,,\ q) = \frac{q^{x^c} - q^{(x+1)^c}}{q^c}\,, \quad x = 1, 2, 3, \ldots\,; c > 0\,; 0 < q < 1, \tag{3.4}$$

em que $q = e^{-(1/\theta)^c}$ e a distribuição denotada por $WeDTZ(q, c)$.

iv. Para $c = 1$ e $k = 1$, essa distribuição se torna a distribuição geométrica truncada em zero (Klugman et al., 2012), isto é:

$$P(X = x \mid q) = \frac{q^x - q^{(x+1)}}{q}\,; \quad x = 1, 2, 3, \ldots\,;\ 0 < q < 1\,. \tag{3.5}$$

em que $q = e^{-(1/\theta)}$ e a distribuição denotada por $GeoTZ(q)$.

v. Tomando $c = 2$, $k = 1$, obtém-se a distribuição Rayleigh discreta truncada em zero com função massa de probabilidade representada por:

$$P(X = x \mid q) = \frac{q^{x^2} - q^{(x+1)^2}}{q^2}\,; \quad x = 1, 2, 3, \ldots\,;\ 0 < q < 1\,. \tag{3.6}$$

em que $q = e^{-(1/\theta)^2}$ e a distribuição denotada por $RaDTZ(q)$.

vi. Com os parâmetros $c = 2$ e $k \longrightarrow \dfrac{k}{2}$, a distribuição se torna a distribuição Rayleigh generalizada discreta truncada em zero, isto é,

$$P(X = x \mid k\,,\theta) = \frac{\Gamma\left(\dfrac{k}{2}\,,\left(\dfrac{x}{\theta}\right)^2,\left(\dfrac{x+1}{\theta}\right)^2\right)}{\Gamma\left(\dfrac{k}{2}\,,\left(\dfrac{1}{\theta}\right)^2\right)}, x = 1,2,3,\ldots\ ;\ k > 0\ ;\ \theta > 0\ . \quad (3.7)$$

Tal distribuição é denotada por $RaGDTZ(k\,,\theta)$.

vii. Para $c = 2$, $k = 3/2$ e $\theta \longrightarrow \sqrt{\theta}$, a distribuição se reduz a Maxwell-Boltzmann discreta truncada em zero, com função massa de probabilidade formulada por:

$$P(X = x \mid \theta) = \frac{\Gamma\left(\dfrac{3}{2}\,,\left(\dfrac{x}{\theta}\right)^2,\left(\dfrac{x+1}{\theta}\right)^2\right)}{\Gamma\left(\dfrac{3}{2}\,,\left(\dfrac{1}{\theta}\right)^2\right)}, x = 1,2,3,\ldots\ ;\ \theta > 0\ ; \quad (3.8)$$

em que a distribuição é denotada por $MaDTZ(\theta)$.

viii. Com $c = 2$ e $k = \dfrac{1}{2}$, essa distribuição se torna a distribuição half-normal discreta truncada em zero, isto é, a função massa de probabilidade é dada por:

$$\begin{aligned} P(X = x \mid \theta) &= \frac{\Gamma\left(\dfrac{1}{2},\left(\dfrac{x}{\theta}\right)^2\right) - \Gamma\left(\dfrac{1}{2},\left(\dfrac{x+1}{\theta}\right)^2\right)}{\Gamma\left(\dfrac{1}{2},\left(\dfrac{1}{\theta}\right)^2\right)} \\ P(X = x \mid a,b) &= \frac{\Phi\left(b\sqrt{2}\right) - \Phi\left(a\sqrt{2}\right)}{\Phi\left((b-a)\sqrt{2}\right)}, \end{aligned} \quad (3.9)$$

em que $a = \dfrac{x}{\theta}$; $b = \dfrac{(x+1)}{\theta}$; $x = 1,2,3,\ldots; b > a > 0;\ \theta > 0$; sendo a distribuição denotada por $HNDTZ(a,b)$.

ix. Para $k \longrightarrow \infty$, $\mu = \log\ \theta + \dfrac{1}{c} log\ k$ e $\sigma = \dfrac{1}{c}\sqrt{k}$, a distribuição reduz a distribuição log-normal discreta truncada em zero, isto é,

$$P(X = x \mid \mu\ ,\ \sigma) = \frac{\Phi\left[\dfrac{\log(x+1)-\mu}{\sigma}\right] - \Phi\left[\dfrac{\log(x)-\mu}{\sigma}\right]}{\Phi\left[\dfrac{-\mu}{\sigma}\right]} \quad (3.10)$$

em que $x = 1,2,3,\ldots;\ \mu > 0\ ;\ \sigma > 0$ e a distribuição denotada por $LNDTZ(\mu\ ,\ \sigma)$.

x. Se o parâmetro for $c = 2$, tal distribuição, reduz a distribuição hidrografic discreta truncada em zero, ou seja,

$$P(X = x \mid k\,,\,\theta) = \frac{\int_x^{x+1} u^{(2k-1)} e^{-(u/\theta)^2} du}{\int_1^{\infty} u^{(2k-1)} e^{-(u/\theta)^2} du} \tag{3.11}$$

em que $x = 1, 2, 3, \ldots ; k > 0$; $\theta > 0$, com a distribuição denotada por $HiDTZ(k\,,\theta)$.

Seja uma amostra aleatória $x_1, x_2, \ldots, x_n$ e a distribuição $GaGDTZ(k, \theta, c)$ como candidatas as estimação, então a função de verossimilhança é dada por

$$L(k, \theta, c \mid \mathbf{x}) = \prod_{i=1}^{n} \frac{\Gamma\left(k, \left(\frac{x}{\theta}\right)^c, \left(\frac{x+1}{\theta}\right)^c\right)}{\Gamma\left(k, \left(\frac{1}{\theta}\right)^c\right)} \tag{3.12}$$

aplicando o logaritmo natural em (3.12) tem-se a função de log verossimilhança

$$l(k, \theta, c \mid \mathbf{x}) = \sum_{i=1}^{n} \log\left[\Gamma\left(k, \left(\frac{x}{\theta}\right)^c, \left(\frac{x+1}{\theta}\right)^c\right)\right] - \sum_{i=1}^{n} \log\left[\Gamma\left(k, \left(\frac{1}{\theta}\right)^c\right)\right] \tag{3.13}$$

Como as equações de verossimilhança não têm solução de forma fechada, há necessidade de um método numérico para encontrar a solução deste sistema não linear.

Capítulo 4

Metodologia

4.1 Determinação dos Dados de Estudo

O banco de dados utilizado para estudo foi obtido a partir dos registros do portal do Instituto Nacional de Meteorologia (INMET) e Agência Nacional da Águas (ANA) para estação de Campo Grande, de Porto São José e Porto Rico do rio Paraná. Ressalta-se que estação de Porto Velho não foi não avaliada por falta de uma série histórica completa com no mínimo 25 anos. Para o estudo foi utilizado as séries históricas do período de 1 de janeiro de 1961 a 31 de dezembro de 2018 para a maioria das estações meteorológicas como segue na Tabela 2.

Tabela 2 – Descrição das séries históricas das estações meteorológicas estudadas.

OMM	Estação	Série histórica		Latitude (°)	Longitude (°)	Altitude (m)
		início	final			
83096	Aracaju	01/01/1961	31/12/2018	-10.95	-37.04	4.72
82397	Fortaleza	01/01/1961	31/12/2018	-3.81	-38.53	26.45
82798	João Pessoa	01/01/1961	31/12/2018	-7.10	-34.86	7.43
82994	Maceió	01/01/1961	31/12/2018	-9.66	-35.70	64.50
82598	Natal	01/01/1961	31/12/2018	-5.91	-35.20	48.60
82994	Recife	01/03/1961	31/12/2018	-8.05	-34.95	10.00
83229	Salvador	01/01/1961	31/12/2018	-13.01	-38.53	51.41
82280	São Luís	01/01/1971	31/12/2018	-2.53	-44.21	50.86
82578	Teresina	01/01/1961	31/12/2018	-5.08	-42.81	74.36
82191	Belém	01/01/1961	31/12/2018	-1.43	-48.43	10.0
82024	Boa Vista	01/01/1972	31/12/2018	2.82	-60.66	83.0
82098	Macapá	01/01/1968	31/12/2018	-0.05	-51.11	14.46
82331	Manaus	01/01/1961	31/12/2018	-3.10	-60.01	61.25
83033	Palmas	01/01/1994	31/12/2018	-10.19	-48.3	280.00
82915	Rio Branco	01/06/1969	31/12/2018	-9.96	-67.80	160.00
83377	Brasília	01/05/1962	31/12/2018	-15.78	-47.92	1159.54
2054014	Campo Grande	01/01/1976	31/12/2018	-20.45	-54.62	562.00
83361	Cuiabá	01/01/1961	31/12/2018	-15.61	-56.10	145.00
83423	Goiânia	01/01/1961	31/12/2018	-16.66	-49.25	741.48

Cont. Tabela 2: Descrição das séries históricas das estações meteorológicas estudadas.

OMM	Estação	Série histórica		Latitude (°)	Longitude (°)	Altitude (m)
		início	final			
83587	Belo Horizonte	01/01/1961	31/12/2018	-19.93	-43.93	915.00
83743	Rio de Janeiro	01/01/1961	31/12/2016	-22.89	-43.18	11.10
83781	São Paulo	01/01/1961	31/12/2018	-23.50	-46.61	792.06
83648	Vitória	01/01/1961	31/12/2018	-20.31	-40.31	36.20
83842	Curitiba	01/01/1961	31/12/2018	-25.43	-49.26	923.50
83897	Florianópolis	01/07/1961	31/12/2018	-27.58	-48.56	1.84
83967	Porto Alegre	01/01/1961	31/12/2018	-30.05	-51.16	46.97
2253002	Porto Rico	01/01/1971	31/12/2018	-22.77	-53.27	240.00
2253016	Porto São José	01/05/1985	31/12/2018	-22.71	-53.17	240.00

4.2 Análise dos Dados

4.2.1 Análise Exploratória

Inicialmente foi feita a análise exploratória dos dados. Para fazer a contagem das sequências de dias chuvosos e as análises das séries históricas anuais, escolheu-se os limiares superior ao volume em milímetros de: 1, 2, 3, 4 e 5. Para as análises das séries históricas mensais das estações do Porto São José e Porto Rico utilizou-se o limiar superior a 5 mm/dia.

A contagem das sequências de dias chuvosos foi realizada por meio da função rle pertencente ao pacote base do software R.

A Figura 5 representa a realização da contagem mensal das sequências de dias chuvosos, a linha vermelha indica o limiar 2, então nessa contagem só fará parte os dias que choveram mais de 2 mm/dias, como segue a descrição da contagem mensal:

- no dia 01/03 não choveu, mas choveu no dia 02/03, porém no dia 03/03 choveu abaixo do limiar, então nesse período, o 02/03 será computado como o dia sem chuva consecutivo, logo o tamanho das sequências é 1 dia;
- não choveu no dia 04/03, mas choveu nos dias 05 e 06/03 mais que o limiar, então o tamanho das sequências é 2 dias;
- passados alguns dias foi chover em 09, 10 e 11/03, mais que o limiar, então o tamanho das sequências é 3 dias;
- no dia 16/03, não choveu antes e depois, logo o tamanho das sequências é 1 dia.
- nos dias 18, 19, 20 e 21/03 ocorreram chuvas acima do limiar, então o tamanho das sequências é 4 dias;
- com chuvas nos dias $26, 27, 28, 29, 30, 31/03$ têm-se que o tamanho das sequências foi 6 dias.

Assim, a contagem sequencial mensal de março ou o tamanho das sequências de dias chuvosos do mês de março ficou: 1, 2, 3, 1, 4, 6. Seguindo o mesmo procedimento em abril ficou: 1, 4, 1, 2, 1.

No entanto, com o objetivo de analisar o tamanho das sequências de dias chuvosos anualmente, a contagem sequencial nesse período exemplificado ficaria: 1, 2, 3, 1, 4, 7, 4, 1, 2, 1.

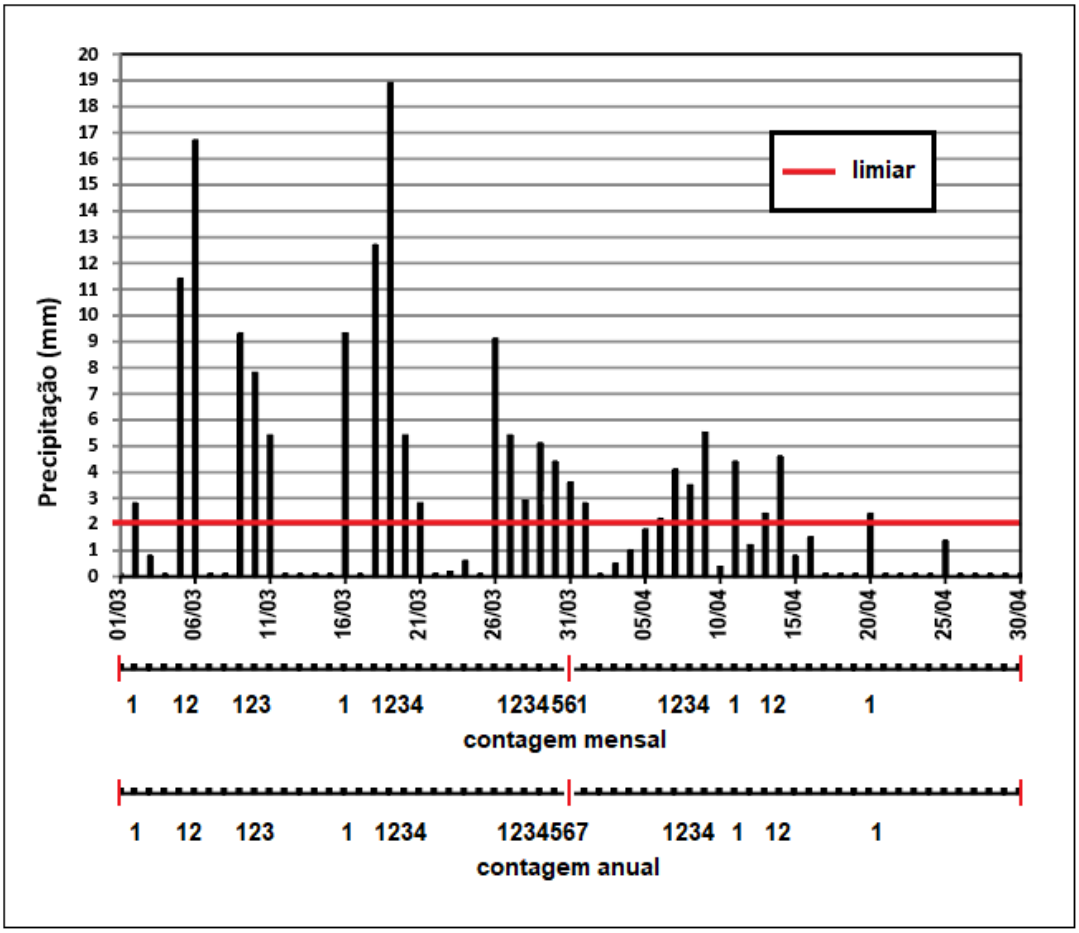

Figura 5 – Representação da contagem do tamanho das sequências dos dias chuvosos.

4.2.2 Teste Wald-Wolfowitz

A independência é uma suposição essencial nas análises das séries, pois nenhuma das observações presentes na amostra pode influenciar a ocorrência, ou a não ocorrência, de qualquer outra observação seguinte (WMO 2009). Segundo Wald e Wolfowitz (1943), dada uma amostra $x_1, x_2, ..., x_n$, de tamanho n, sob as hipóteses formuladas por

H_0: As observações são independentes

H_1: As observações não são independentes .

A estatística do teste de Wald-Wolfowitz, para n grande, é dada por

$$Z = \frac{R - E(R)}{Var(R)} \tag{4.1}$$

em que:

$$R = \sum_{i=1}^{n-1} x_1 x_{i+1} + x_n x_1 \ ;$$
$$E(R) = \frac{-s_2}{(n-1)} \ \text{ e } \ Var(R) = \frac{(s_2^2 - s_4)}{(n-1)} + \frac{(s_2^2 - 2s_4)}{(n-1)(n-2)} - \frac{s_2^2}{(n-1)^2} \ .$$

O teste de independência foi realizado no software R função *ww.test* via biblioteca *trend* (Pohlert, 2018).

4.2.3 Teste Mann-Kendall

O teste de Mann-Kendall é utilizado para testar se as observações de uma série são independentes e identicamente distribuída, assim, sejam as observações $x_1, x_2, ..., x_n$ de uma série temporal, sob as hipóteses dada por (Pohlert, 2018):

H_0: As séries são independentes e identicamente distribuídas, isto é, não há tendência.

H_1: As séries não são independentes e identicamente distribuídas, isto é, há tendência.

A estatística do teste é dada por:

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} sign(x_j - x_k) \tag{4.2}$$

em que

$$sign(x) = \begin{cases} 1, & \text{se } x > 0 \\ 0, & \text{se } x = 0 \\ -1, & \text{se } x < 0 \end{cases}$$

Para grandes amostras $(n > 30)$ a estatística do teste é calculada por:

$$Z = \begin{cases} \frac{S-1}{\sigma}, & \text{se } S > 0 \\ 0, & \text{se } S = 0 \\ \frac{S+1}{\sigma}, & \text{se } S < 0 \end{cases} , \tag{4.3}$$

sendo, S é normalmente distribuída, ou seja, $S \sim N(\mu, \sigma^2)$ então, $\mu = 0$ e

$$\sigma^2 = \frac{n(n-1)(2n+5) - \sum_{j=1}^{P} t_j(t_j - 1)(2t_j + 5)}{18},$$

em que n é o número de observações , P é o número de grupos com observações iguais e t_j é o numero de observações iguais no grupo j.

Assim, hipótese nula (H_0) não é rejeitada se o p-valor for superior ao um dado nível de significância, isto é, as observações da série são independentes e identicamente distribuída.

O teste Mann-Kendall foi realizado no software R com função *mk.test* via biblioteca *trend* (Pohlert, 2018).

4.2.4 Estimação de Máxima Verossimilhança

Os parâmetros das distribuições estudadas foram estimados por meio do método da máxima verossimilhança. A estimação foi executada pela função *fitdist* da biblioteca *fitdistplus* do software R (Delignette-Muller e Dutang, 2015). As estimativas intervalares dos parâmetros foram determinadas pelo método Delta (Davison, 2011).

4.2.5 Teste de Aderência

Para verificar se a distribuição de probabilidade utilizada ajusta-se a série de dados disponível, podem-se aplicar alguns testes não paramétricos, tais como o teste Qui-quadrado Pearson. O teste é simbolizado por χ^2, mede a adequabilidade do ajuste da distribuição, ou seja, o quanto a frequência observada está próxima do valor esperado. As hipóteses são formuladas por

H_0: A distribuição de probabilidade proposta é adequada
H_1: A distribuição de probabilidade proposta não é adequado

A estatística do teste descreve a qualidade do ajuste dos dados ao modelo é dada por

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{O_i - E_i}{E_i} \right)^2 \tag{4.4}$$

em que:

O_i: valor observado do período chuvoso ou seco
E_i: valor esperado do período chuvoso ou seco
$i = 1, 2, 3, \ldots, n$
n: número de classes
$\nu = n - t - 1$
t: número de parâmetros estimados de cada distribuição.

Deve-se notar que as distribuições com um valor χ^2 mais baixo e uma maior probabilidade produzem um melhor ajuste. Por outro lado, se nenhum dos as distribuições se encaixam na distribuição dos períodos chuvosos e secos em uma determinada estação pluviométrica, estatística mínima χ^2 será considerada para o seleção da distribuição mais apropriada (Deni et al., 2008).

Para verificar se a distribuição de probabilidade utilizada ajusta-se a série de dados estudada a qualidade do ajuste foi considerando nível de significância de 1%. Tal teste foi realizado pela função *gofstat* do pacote *fitdistplus* do software estatístico R (Delignette-Muller e Dutang, 2015)

4.2.6 Critérios de Informação

Os testes de ajuste de modelo são aplicados para avaliar a adequação de diferentes distribuições de probabilidade, isto é, esses testes estabelecem quais distribuições são, em geral, as mais apropriadas para modelagem de determinado fenômeno.

4.2.6.1 Critérios de Informação de Akaike

Para discriminar os modelos ajustados, Akaike (1974) introduziu um critério de informação chamado AIC, que significa critério de informação de Akaike. Pode ser adaptado a muitas situações diferentes e consiste em minimizar uma medida de informação. O critério de informação é definido da seguinte forma:

$$AIC = -2 \log L(\widehat{\theta}) + 2p \tag{4.5}$$

em que:

$L(\widehat{\theta})$: função de verossimilhança
p: número de parâmetros.

O modelo que melhor explica os dados é aquele com o menor critério de informação de Akaike. Além de selecionar o ajuste mais adequado, o critério penaliza a adição de parâmetros, levado em consideração o princípio da parcimônia, isto é, selecionar modelos mais simples (Akaike, 1974).

4.2.6.2 Critério de Informação Bayesiano

Uma extensão bayesiana do conceito mínimo de critério de informação de Akaike é o critério de informação bayesiano chamado BIC. Schwarz et al. (1978) propôs o critério de informação definido da seguinte forma:

$$BIC = -2 \log L(\widehat{\theta}) + p \log(n) \tag{4.6}$$

em que:

$L(\widehat{\theta})$: função de verossimilhança
p: número de parâmetros
n: tamanho da amostra.

O critério de informação bayesiano também é um critério de parcimônia. De todos os modelos, aquele com o menor critério de informação bayesiana é considerado o melhor.

Os critérios de informação de AIC e BIC foram determinados pela função *gofstat* do também pelo pacote *fitdistplus* (Delignette-Muller e Dutang, 2015).

Capítulo 5

Resultados e Discussões

5.1 Análises das Séries Anuais

Para o estudo das análises das séries históricas anuais foi utilizado a distribuição gama generalizada truncada em zero para descrever o tamanho das sequências de dias chuvosos para todas as capitais do Brasil.

5.1.1 Região Nordeste

A análise exploratória do tamanho das sequências de dias chuvosos para as séries históricas das estações das capitais da Região Nordeste indicou por meio do coeficiente de variação uma grande variabilidade no tamanho das sequências de dias chuvosos independente do limiar utilizado e estações, assim como a moda foi de um dia chuvoso para todas configurações estudadas (Tab. 3).

Com limiar 1 mm/dia de chuva, constatou-se que a maior sequência do tamanho de dias chuvosos foi de 32 dias ocorrido na estação de Recife, ao aumentar o limiar para 4 mm/dia de chuva foi possível constatar que a sequência longa de chuva caiu para 16 dias, um indicativo que as sequências consecutivas de mais 16 dias foram de chuvas fracas. Na estação de Maceió, também teve ocorrência como esta porém com 30 dias consecutivos de chuvas. Notou-se que com limiar 5 mm/dia de chuva, a maior sequência do tamanho de dias chuvosos foi de 22 dias ocorrido na estação de Fortaleza. Ressalta-se, também, que a mudança de limiar não influenciou no tamanho de dias chuvosos da estação de Aracaju (Tab. 3; Figs. 6, 7 e 8).

Salienta-se que a validade das suposições de que as séries históricas no tamanho das sequências de dias chuvosos das capitais da Região Nordeste, com nível de significância de 1%, mostraram-se independentes e identicamente distribuídas, segundo os testes Mann-Kendall e Wald-Wolfowitz.

Tabela 3 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas anuais de 1961 a 2018 das estações meteorológicas das capitais da Região do Nordeste.

Capital	Limiar	n	Mo	Q_1	Md	$\overline{x}$	Q_3	Max	SD	CV (%)
Aracaju	1	5887	1	1	2	2.29	3	15	1.94	84.48
	2	4849	1	1	1	1.99	2	15	1.53	76.72
	3	4152	1	1	1	1.84	2	15	1.37	74.42
	4	3550	1	1	1	1.75	2	15	1.27	72.80
	5	3101	1	1	1	1.69	2	15	1.20	71.17
Fortaleza	1	5482	1	1	1	2.40	3	29	2.48	103.22
	2	4555	1	1	1	2.21	3	28	2.11	95.65
	3	3959	1	1	1	2.05	2	24	1.79	87.30
	4	3566	1	1	1	1.98	2	24	1.69	85.30
	5	3237	1	1	1	1.90	2	22	1.56	81.91
João Pessoa	1	7125	1	1	2	2.41	3	25	2.24	93.02
	2	5955	1	1	1	2.15	3	25	1.82	84.63
	3	5149	1	1	1	2.03	2	15	1.57	77.55
	4	4588	1	1	1	1.91	2	12	1.39	72.88
	5	4059	1	1	1	1.80	2	12	1.26	70.30
Maceió	1	6758	1	1	2	2.98	4	30	3.14	105.16
	2	5629	1	1	2	2.54	3	22	2.40	94.52
	3	4851	1	1	2	2.25	3	22	1.94	86.18
	4	4252	1	1	1	2.10	3	16	1.76	84.06
	5	3818	1	1	1	1.98	2	16	1.60	80.78
Natal	1	5203	1	1	2	2.20	3	20	1.88	85.60
	2	4296	1	1	1	2.00	2	16	1.54	77.21
	3	3723	1	1	1	1.87	2	12	1.38	73.99
	4	3283	1	1	1	1.76	2	12	1.25	71.03
	5	2916	1	1	1	1.69	2	12	1.18	69.76
Recife	1	9462	1	1	2	2.71	3	32	2.63	96.93
	2	8015	1	1	2	2.37	3	25	2.13	89.80
	3	7032	1	1	1	2.17	3	25	1.85	85.48
	4	6217	1	1	1	2.01	2	16	1.62	80.85
	5	5560	1	1	1	1.92	2	16	1.50	78.52
Salvador	1	7145	1	1	2	2.57	3	18	2.25	87.48
	2	6071	1	1	2	2.25	3	15	1.81	80.68
	3	5204	1	1	1	2.03	3	14	1.56	76.63
	4	4579	1	1	1	1.89	2	13	1.37	72.69
	5	4080	1	1	1	1.79	2	13	1.29	71.97

Legenda: Número de dias chuvosos (n); moda (Mo); $1^{\underline{o}}$ quartil (Q_1); mediana (Md); média ($\overline{x}$); $3^{\underline{o}}$ quartil (Q_3); valor máximo (Max); desvio padrão (sd); coeficiente de variação (cv).

Cont. Tabela 3 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas anuais de 1961 a 2018 das estações meteorológicas das capitais da Região do Nordeste.

Capital	Limiar	n	Mo	Q_1	Md	$\overline{x}$	Q_3	Max	SD	CV (%)
São Luís	1	5837	1	1	2	2.69	3	28	2.65	98.74
	2	5218	1	1	2	2.40	3	23	2.15	89.86
	3	4753	1	1	2	2.18	3	23	1.86	85.02
	4	4417	1	1	1	2.07	3	23	1.69	81.53
	5	4086	1	1	1	1.95	2	15	1.51	77.51
Teresina	1	3307	1	1	1	1.94	2	15	1.67	86.08
	2	2902	1	1	1	1.81	2	15	1.46	80.82
	3	2574	1	1	1	1.71	2	11	1.28	75.01
	4	2332	1	1	1	1.61	2	11	1.15	71.35
	5	2164	1	1	1	1.53	2	10	1.01	65.82

Legenda: Número de dias chuvosos (n); moda (Mo); $1^{\underline{o}}$ quartil (Q_1); mediana (Md); média ($\overline{x}$); $3^{\underline{o}}$ quartil (Q_3); valor máximo (Max); desvio padrão (sd); coeficiente de variação (cv).

Tabela 4 – Teste qui-quadrado de aderência para a distribuição $GaGDTZ(k, \theta, c)$ ajustada as séries históricas das estações meteorológicas das capitais da Região Nordeste de 1961 a 2018.

Estação	Estatística do teste	Limiar 1	2	3	4	5
Aracaju	χ^2	14.3067	12.6899	6.2803	6.0957	5.0053
	p valor	0.1594	0.1229	0.5074	0.4125	0.5431
Fortaleza	χ^2	23.0280	11.1743	11.7673	7.3858	10.2150
	p valor	0.1483	0.6722	0.3813	0.7670	0.4218
João Pessoa	χ^2	5.0192	7.5619	2.1799	4.1007	0.4033
	p valor	0.9747	0.5788	0.9492	0.6630	0.9988
Maceió	χ^2	16.2285	8.4441	9.4406	9.0251	9.0304
	p valor	0.5766	0.8135	0.3976	0.4349	0.2504
Natal	χ^2	7.2228	4.5090	3.4372	2.5109	5.1726
	p valor	0.7807	0.7196	0.7522	0.8672	0.5218
Recife	χ^2	13.6071	9.9903	4.6984	5.0177	11.0725
	p valor	0.6946	0.6168	0.9448	0.8327	0.2707
Salvador	χ^2	4.7536	4.9817	6.0076	2.3484	1.22302
	p valor	0.9657	0.8923	0.6463	0.3090	0.54253
São Luís	χ^2	16.0747	7.1542	9.3424	12.2160	7.0078
	p valor	0.3088	0.7687	0.4999	0.2014	0.4280
Teresina	χ^2	11.7726	9.4139	11.1204	12.2962	2.8151
	p valor	0.1616	0.2242	0.1334	0.0556	0.5892

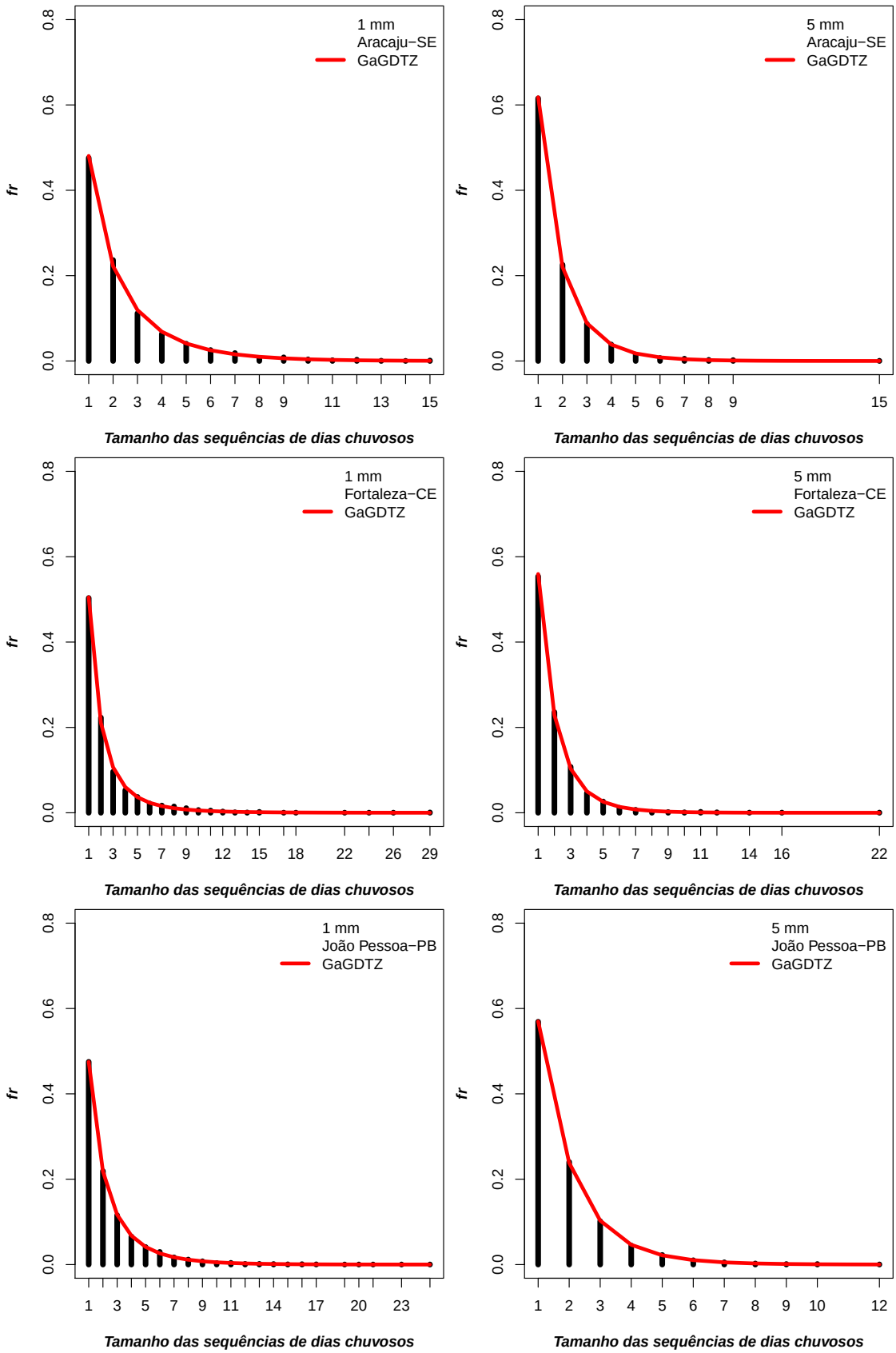

Figura 6 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Nordeste de 1961 a 2018.

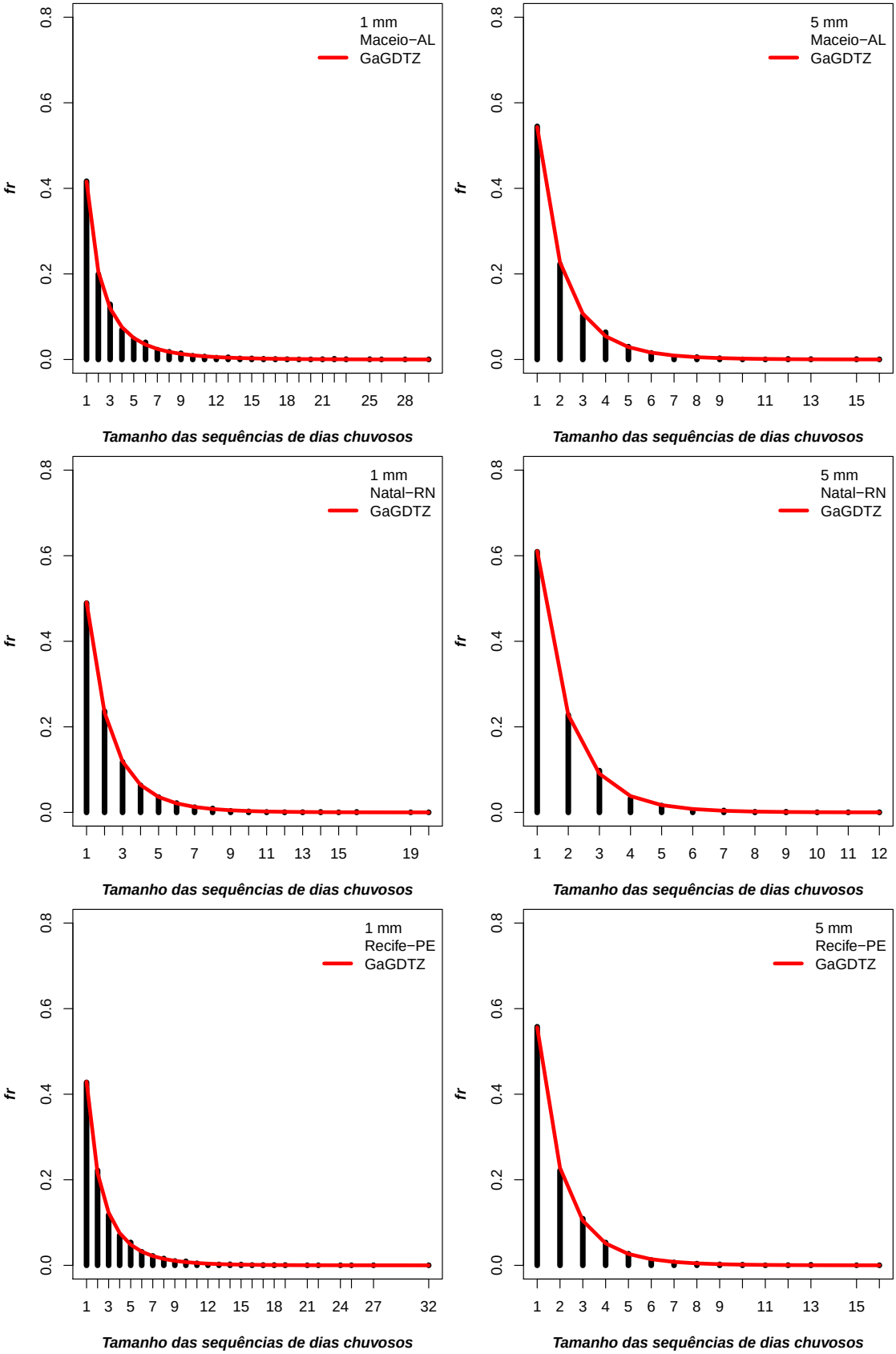

Figura 7 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Nordeste 1961 a 2018.

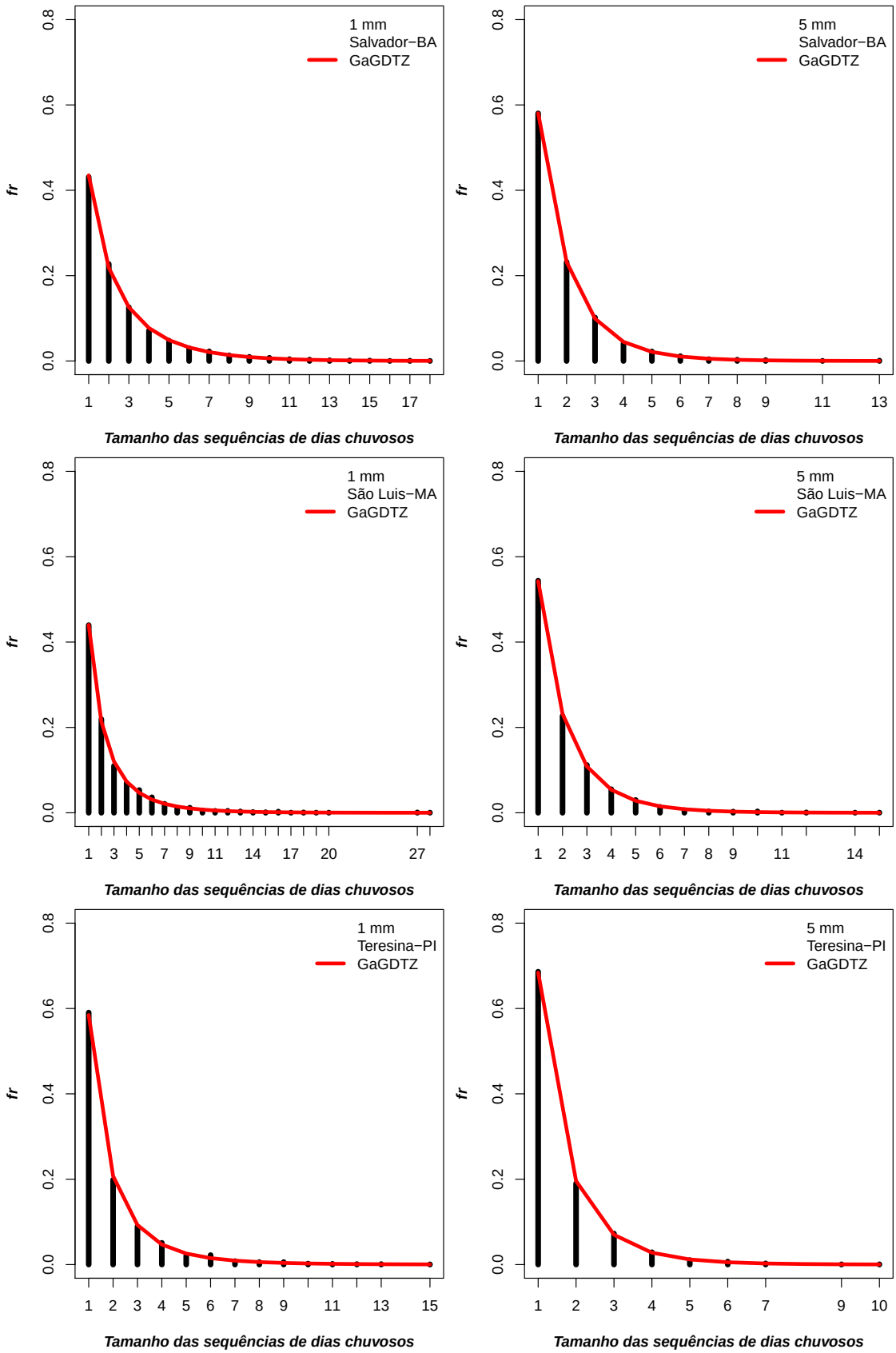

Figura 8 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Nordeste de 1961 a 2018.

Com o teste qui-quadrado (χ^2) de aderência foi possível constatar que o tamanho das sequências de dias chuvosos apresentaram-se adequados no ajuste da distribuição $GaDGTZ(k, \theta, c)$ em todas as séries históricas das capitais da Região do Nordeste, com nível de significância de 5% (Tabela 4) .

As formas das curvas da distribuição $GaDGTZ(k, \theta, c)$ ajustada sobre os valores observados do tamanho das sequências de dias chuvosos, apresentadas nas Figuras 6 , 7 e 8, evidenciam os resultados obtidos pelo teste de adequabilidade dos dados. Devido a extensa amplitude do tamanho da sequências, para melhor visualização dos gráficos foram representados apenas as séries histórias com limiar 1 e 5 mm/dia de chuva.

A Tabela 5 apresenta as estimativas intervalares dos parâmetros em relação a cada limiar, mas para verificar a possibilidade de algum valor da estimativa do parâmetro poder reduzir a distribuição $GaGDTZ\ (k, \theta, c)$ em um dos seus casos particulares, tomou-se como exemplo os dados com limiar 5 mm/dia de chuva.

Assim, descrevendo a estação de Aracaju, tem-se que o intervalo de confiança (IC) de 95% do parâmetro c não contém o valor 2, sugerindo que a distribuição estudada não reduzirá nas distribuições discretas truncadas em zero: Rayleigh, Rayleigh generalizada, hidrografic, Maxwell e half-normal, mas contém o 1 nas estimativas dos parâmetros k, θ, c indicando que a distribuição $GaGDTZ(k, \theta, c)$ pode reduzir nas distribuições discretas truncadas em zero: Gama $(GaDTZ(k, \theta))$, Gama com 1 parâmetro $(GaDTZ(k))$, Geométrica $(GeoTZ(q))$, Weibull $(WeDTZ(q, c))$ (Tab. 5). No entanto, as distribuições GaDTZ(k) e GeoTZ(q) não tiveram ajustes adequados, possivelmente, pela amplitude do tamanho das sequências de chuvas ser extensa e pelas frequências dos valores esperados dessas duas distribuições divergirem dos valores observados. Acrescenta-se ainda que a estimativa intervalar do parâmetro k parece ser longa, um indicativo que a distribuição $GaGDTZ(k, \theta, c)$ pode reduzir na distribuição log-normal discreta truncada em zero $(LNDTZ(\mu, \sigma))$.

Do mesmo modo, as análises mostraram que em todas as estações das capitais da Região Nordeste podem-se notar uma amplitude longa e o valor 1 contido nas estimativas intervalares do parâmetro k. Diante disso, testou-se o desempenho dos ajustes das distribuições reduzidas em relação o tamanho das sequências de dias chuvosos, salientando-se que a Tabela 6 apresenta o teste χ^2 de aderência para pelo menos uma estação com ajuste adequado então, constatou-se pelo teste que as distribuições $WeDTZ(q, c)$ e $LNDTZ(\mu, \sigma)$ foram adequadas para todas as capitais.

Com a ajuda dos critérios de informações AIC e BIC constatou-se que a distribuição $WeDTZ(q, c)$ apresentou melhor desempenho nos ajustes do tamanho das sequências de dias chuvosos para todas as estações, exceto em Fortaleza que AIC foi menor para a distribuição $LNDTZ(\mu, \sigma)$ (Tabela 6).

Tabela 5 – Estimativa intervalar dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada do tamanho das sequências de dias chuvosos das capitais da Região Nordeste de 1961 a 2018.

Estação	Parâmetros	Limiar 1		Limiar 2		Limiar 3		Limiar 4		Limiar 5	
		LI	LS	LI	LS	LI	LS	LI	LS	LI	LS
Aracaju	k	0.004	9.480	0.003	9.431	0.418	8.580	0.243	9.086	0.210	8.799
	θ	0.903	9.872	0.996	7.033	0.035	4.181	0.042	5.001	0.049	4.637
	c	0.629	1.726	0.721	1.743	0.424	1.063	0.431	1.141	0.443	1.152
Fortaleza	k	0.069	10.447	1.055	9.109	0.537	8.232	1.739	8.845	1.348	8.699
	θ	0.022	10.743	0.003	1.395	0.013	3.012	0.005	0.563	0.009	1.003
	c	0.320	0.898	0.315	0.639	0.357	0.803	0.355	0.620	0.382	0.716
João Pessoa	k	0.438	10.368	0.364	10.027	0.341	9.755	0.251	9.783	0.305	9.345
	θ	0.009	7.511	0.018	7.076	0.035	6.933	0.058	7.491	0.051	5.612
	c	0.321	0.895	0.359	1.017	0.405	1.177	0.438	1.355	0.449	1.266
Maceió	k	0.002	12.052	0.016	8.229	0.007	8.157	0.023	7.134	0.204	7.332
	θ	1.047	13.091	0.414	12.631	0.692	9.016	0.347	7.992	0.067	5.830
	c	0.556	1.340	0.492	1.416	0.594	1.527	0.529	1.374	0.433	1.120
Natal	k	0.855	10.614	0.002	8.174	0.115	5.709	0.310	4.866	0.173	5.504
	θ	0.008	3.483	1.291	6.173	0.212	5.609	0.157	3.108	0.169	3.865
	c	0.349	0.846	0.807	1.657	0.539	1.354	0.551	1.132	0.555	1.258
Recife	k	0.417	11.059	0.024	9.083	0.006	9.275	0.027	6.041	0.320	6.760
	θ	0.009	11.250	0.282	12.845	0.601	8.307	0.389	6.789	0.058	4.105
	c	0.310	0.941	0.471	1.480	0.584	1.458	0.561	1.372	0.441	1.032
Salvador	k	0.004	9.879	0.103	9.466	0.059	9.547	0.041	9.254	0.186	9.593
	θ	1.114	13.001	0.129	13.135	0.187	10.995	0.255	9.065	0.061	6.655
	c	0.622	1.955	0.440	1.604	0.483	1.699	0.532	1.768	0.446	1.321
São Luís	k	0.062	7.209	0.208	7.569	0.298	7.058	0.149	7.241	0.257	7.421
	θ	0.165	13.749	0.066	9.151	0.069	6.066	0.122	7.491	0.070	5.538
	c	0.416	1.236	0.401	1.134	0.432	1.099	0.469	1.288	0.447	1.161
Teresina	k	$<$0.000	8.148	$<$0.000	8.538	$<$0.000	6.959	$<$0.000	6.155	$<$0.000	6.498
	θ	0.769	2.361	0.757	2.287	0.705	2.243	0.615	2.063	0.750	2.119
	c	0.582	0.925	0.609	0.985	0.622	1.044	0.624	1.064	0.702	1.189

Tabela 6 – Teste qui-quadrado de aderência e critérios de informação para os modelos ajustados ao tamanho das sequências de dias chuvosos das capitais da Região Nordeste de 1961 a 2018 - limiar 5.

Estação	Estatística do teste	Distribuição GaDTZ	WeDTZ	LNDTZ	GeoTZ
Aracaju	χ^2	3.6390	2.6340	1.8138	24.7750
	p valor	0.7253	0.8531	0.8742	0.0008
Fortaleza	χ^2	11.8218	8.3382	3.6952	64.80809
	p valor	0.1593	0.4011	0.8141	<0.0000
João Pessoa	χ^2	1.7007	1.0572	1.5651	9.2555
	p valor	0.9745	0.9938	0.9550	0.3212
Maceió	χ^2	10.7745	9.2130	11.2255	75.1466
	p valor	0.2148	0.3246	0.1290	<0.0000
Natal	χ^2	7.1131	5.7257	3.5291	22.2980
	p valor	0.3105	0.4546	0.6189	0.0022
Recife	χ^2	4.1632	3.0610	6.8174	62.1281
	p valor	0.9003	0.9618	0.5564	<0.0000
Salvador	χ^2	2.6681	2.0197	2.9766	17.1790
	p valor	0.8491	0.9178	0.7035	0.0162
São Luís	χ^2	7.8055	7.0957	9.5559	37.5565
	p valor	0.4526	0.5263	0.2151	<0.0000
Teresina	χ^2	18.8598	1.7998	3.3862	23.0062
	p valor	0.0008	0.7725	0.3358	0.0003
Estação	**Critério de Informação**	**Distribuição GaDTZ**	**WeDTZ**	**LNDTZ**	**GeoTZ**
Aracaju	AIC	4181.033	4179.871	4181.191	4195.156
	BIC	4192.064	4190.902	4197.737	4200.671
Fortaleza	AIC	4457.901	4453.768	4449.924	4481.223
	BIC	4468.777	4464.645	4466.239	4486.662
João Pessoa	AIC	5573.601	5573.178	5577.155	5577.406
	BIC	5585.046	5584.622	5594.322	5583.128
Maceió	AIC	5264.873	5263.964	5268.968	5294.453
	BIC	5276.003	5275.094	5285.662	5300.018
Natal	AIC	3937.149	3935.954	3936.064	3943.855
	BIC	3948.058	3946.863	3952.428	3949.310
Recife	AIC	7661.033	7659.628	7665.080	7698.909
	BIC	7672.980	7671.575	7683.000	7704.882
Salvador	AIC	5588.205	5587.451	5590.906	5599.091
	BIC	5599.673	5598.920	5608.108	5604.825
São Luis	AIC	5644.642	5643.956	5649.288	5663.437
	BIC	5655.342	5654.783	5666.205	5669.296
Teresina	AIC	2790.943	2776.665	2780.883	2793.606
	BIC	2801.453	2787.175	2796.647	2798.861

5.1.2 Região Norte

A Região Norte é caracterizada por ter um clima úmido e classificada segundo Köppen como clima tropical (A). Observando a Tabela 7, pode-se constatar pelo coeficiente de variação, uma grande variabilidade no tamanho das sequências de dias chuvosos das estações das capitais da região, deste modo ressalta-se que a moda foi de um dia chuvoso, enquanto que, o terceiro quartil indicou que o maior tamanho das sequências de dias chuvosos foi na estação de Belém com limiar 1 mm/dia de chuva, isto é, em 25% dos dias choveram mais que 4 dias consecutivos.

Tabela 7 – Análise descritiva do tamanho das sequências de dias chuvosos da série histórica de 1961 a 2018 das estações meteorológicas das capitais da Região do Norte.

Capital	Limiar	n	Mo	Q_1	Md	$\overline{x}$	Q_3	Max	SD	CV (%)
Belém	1	12000	1	1	2	3.18	4	38	3.58	112.55
	2	10777	1	1	2	2.70	3	26	2.71	100.58
	3	9844	1	1	2	2.40	3	22	2.24	93.03
	4	9031	1	1	1	2.20	3	17	1.96	89.29
	5	8273	1	1	1	2.03	2	15	1.72	84.79
Boa Vista	1	4432	1	1	1	2.23	3	20	2.16	96.83
	2	3880	1	1	1	2.04	2	20	1.81	71.76
	3	3423	1	1	1	1.88	2	14	1.50	79.63
	4	3113	1	1	1	1.79	2	14	1.35	75.35
	5	2830	1	1	1	1.73	2	14	1.26	72.67
Macapá	1	8036	1	1	2	2.76	3	27	2.55	92.45
	2	7190	1	1	2	2.44	3	22	2.11	86.44
	3	6584	1	1	2	2.25	3	15	1.85	82.55
	4	6055	1	1	1	2.09	3	13	1.63	77.99
	5	5565	1	1	1	1.96	2	11	1.47	75.09
Manaus	1	8105	1	1	2	2.18	3	20	1.87	85.99
	2	7190	1	1	1	1.94	2	19	1.51	77.75
	3	6509	1	1	1	1.78	2	16	1.29	72.73
	4	5938	1	1	1	1.68	2	16	1.15	68.53
	5	5444	1	1	1	1.61	2	16	1.06	65.73
Palmas	1	2790	1	1	2	2.18	3	22	1.77	80.88
	2	2495	1	1	1	1.96	2	14	1.48	75.17
	3	2243	1	1	1	1.80	2	14	1.32	73.43
	4	2043	1	1	1	1.68	2	11	1.18	70.38
	5	1895	1	1	1	1.60	2	10	1.02	64.02
Rio Branco	1	5862	1	1	1	1.94	2	15	1.47	75.46
	2	5119	1	1	1	1.74	2	13	1.22	70.25
	3	4626	1	1	1	1.62	2	10	1.07	65.84
	4	4208	1	1	1	1.53	2	10	0.96	62.78
	5	3867	1	1	1	1.47	2	10	0.89	60.14

Legenda: Número de dias chuvosos (n); moda (Mo); $1^{\underline{o}}$ quartil (Q_1); mediana (Md); média ($\overline{x}$); $3^{\underline{o}}$ quartil (Q_3); valor máximo (Max); desvio padrão (sd); coeficiente de variação (cv).

Avaliando as maiores sequências dos tamanhos de dias chuvosos verificou que na estação de Belém já ocorreu até 38 dias consecutivos de chuvas, mas pode-se notar que essas sequências de chuvas foram fracas, pois quando mudou o limiar para 5 mm/dia de chuvas o tamanho máximo reduziu para 15 dias consecutivos de chuvas. A estação de Rio Branco mostrou dispor das menores sequências dos tamanhos de dias chuvosos (Tab. 7 e Figs. 9 e 10).

Para realizar as estimações intervalares dos parâmetros e verificar a adequação da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos testou-se a validade das suposições de que as séries históricas do tamanho das sequências das capitais da Região Norte seguem independentes e identicamente distribuídas, com nível de significância de 1%, segundo os testes Wald-Wolfowitz e Mann-Kendall tais suposições não foram violadas.

Observando as Figuras 9 e 10, pode-se perceber que a distribuição $GaGDTZ\ (k, \theta, c)$, visualmente, se adequou ao tamanho das sequências de dias chuvosos independente do limiar e das estações avaliadas.

O teste de aderência confirma o observado nas Figuras 9 e 10, pois com nível de significância de 1%, H_o não foi rejeitada para nenhuma das estações com os respectivos limiares, ou seja, a distribuição $GaGDTZ(k, \theta, c)$ apresentou-se bem ajustada ao tamanho das sequências de dias chuvosos para todas as condições propostas no estudo (Tab. 8).

Avaliando as estimativas intervalares do parâmetro c, na Tabela 9, constatou-se que o valor 2 não está contido nos intervalos do parâmetros c, mas o valor 1 está contido nos intervalos dos parâmetros k, θ, c das estações avaliadas, exceto a de Manaus, dessa forma sugerindo que a distribuição $GaGDTZ(k, \theta, c)$ pode reduzir em seus casos particulares.

Tabela 8 – Teste qui-quadrado de aderência para a distribuição $GaGDTZ(k, \theta, c)$ ajustada as séries históricas das estações meteorológicas das capitais da Região Norte de 1961 a 2018.

Estação	Estatística do teste	Limiar				
		1	2	3	4	5
Belém	χ^2	21.3854	13.1217	21.2318	10.4490	7.8161
	p valor	0.3747	0.5169	0.0684	0.4905	0.6467
Boa Vista	χ^2	14.2045	18.6912	5.8465	11.1352	5.1136
	p valor	0.2878	0.0279	0.5577	0.0842	0.5293
Macapá	χ^2	13.3107	11.9292	7.9027	12.2734	5.7506
	p valor	0.2735	0.2898	0.4430	0.1394	0.5692
Manaus	χ^2	14.2169	7.1977	4.3322	2.8009	1.8758
	p valor	0.1633	0.5154	0.6318	0.7306	0.7585
Palmas	χ^2	3.3687	8.3623	8.5617	7.3055	2.8955
	p valor	0.8489	0.2127	0.1278	0.1205	0.5754
Rio Branco	χ^2	4.5457	1.7084	6.4565	4.3578	5.0793
	p valor	0.8048	0.9444	0.1675	0.3597	0.1660

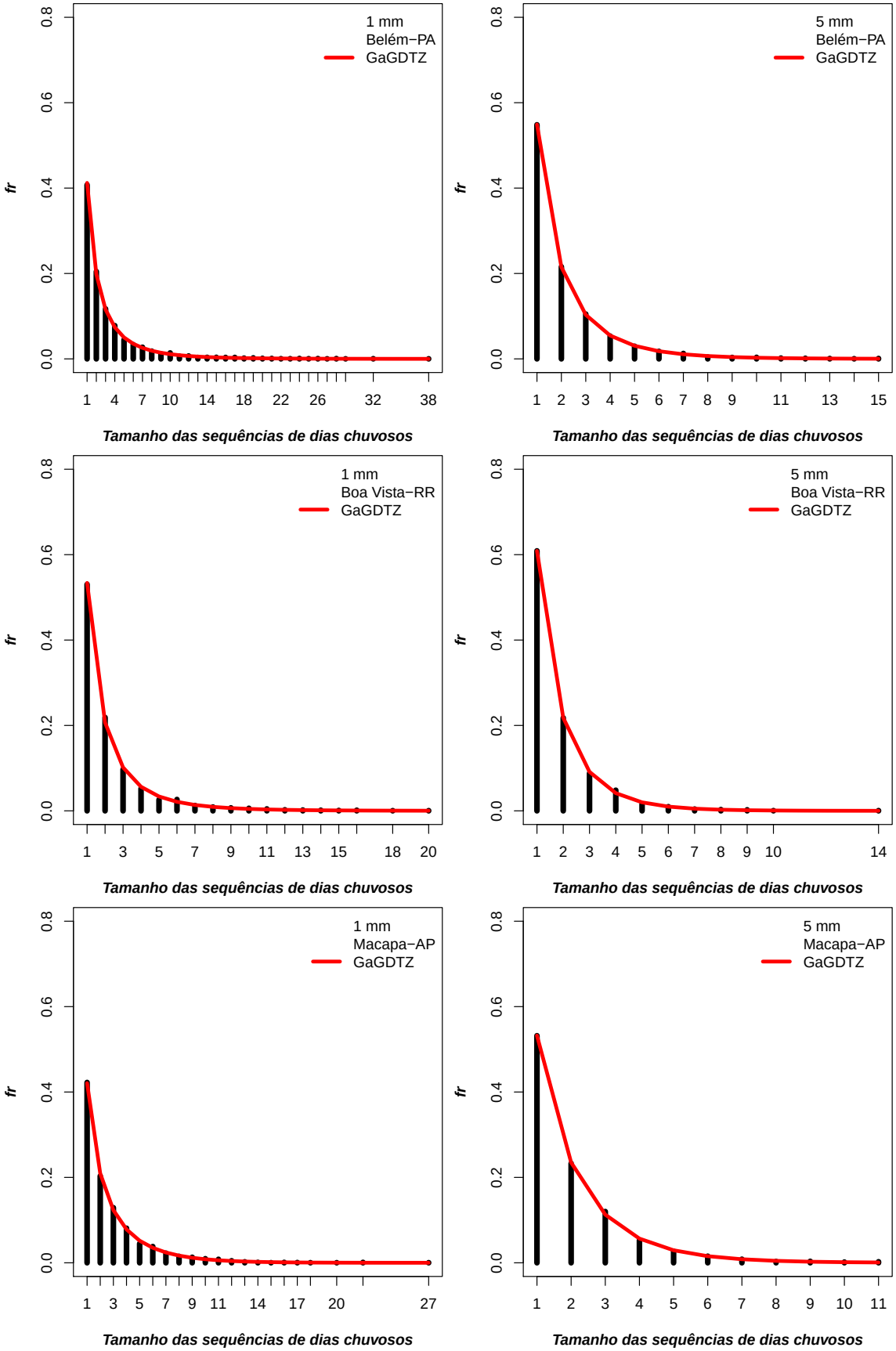

Figura 9 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Norte de 1961 a 2018.

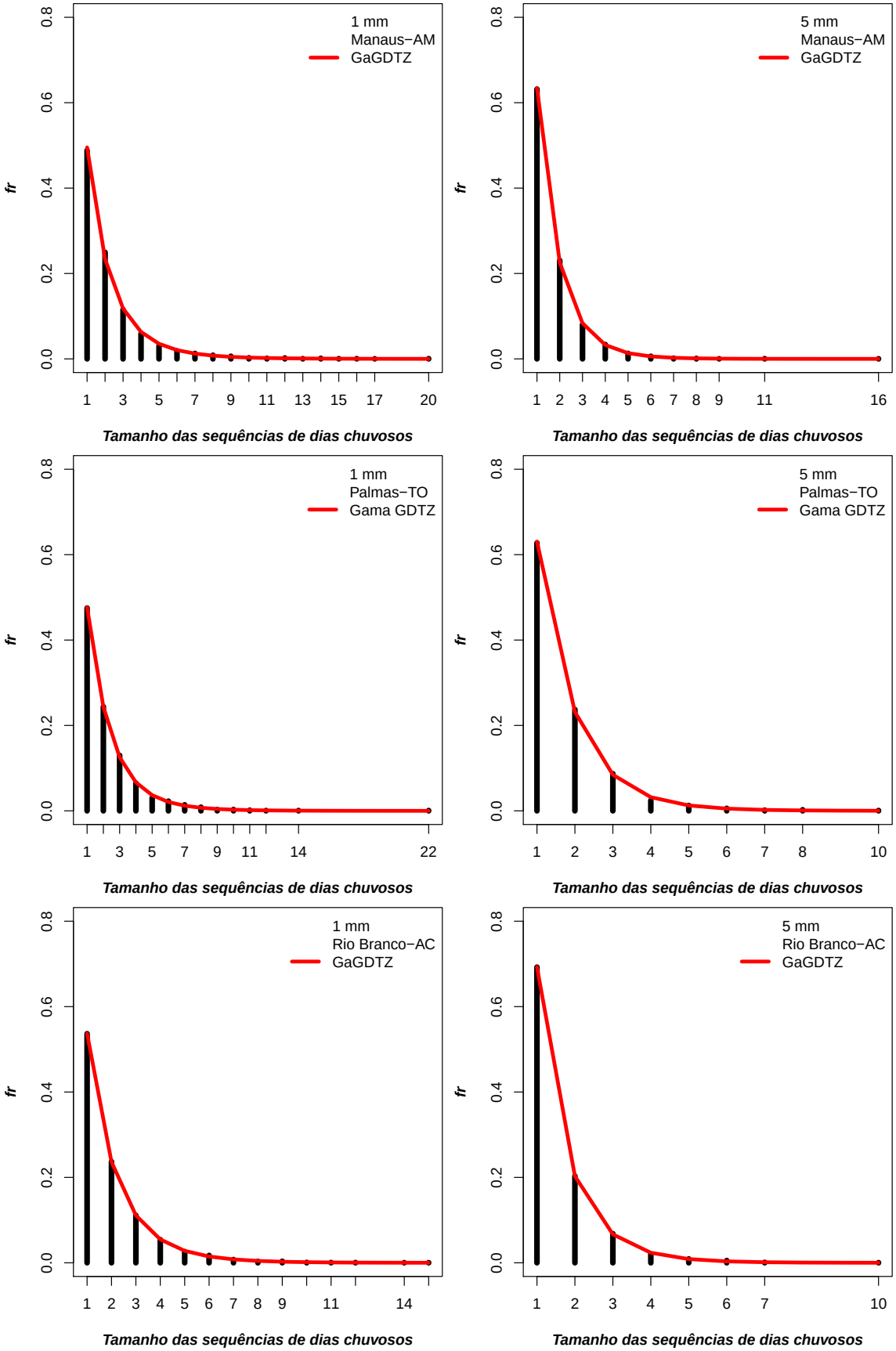

Figura 10 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Norte de 1961 a 2018.

Tabela 9 – Estimação intervalar dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada do tamanho das sequências de dias chuvosos das capitais da Região Norte de 1961 a 2018.

Estação	Parâmetros	Limiar 1		Limiar 2		Limiar 3		Limiar 4		Limiar 5	
		LI	LS	LI	LS	LI	LS	LI	LS	LI	LS
Belém	k	0.047	7.171	0.038	6.792	0.008	6.359	0.072	5.687	0.016	5.780
	θ	0.143	14.338	0.224	12.496	0.616	8.066	0.153	6.785	0.391	5.442
	c	0.375	1.074	0.431	1.180	0.548	1.234	0.448	1.078	0.549	1.154
Boa Vista	k	0.004	11.444	0.001	11.239	0.004	8.636	0.002	8.416	0.021	8.856
	θ	0.242	5.472	0.528	4.089	0.442	4.846	0.643	4.089	0.254	5.706
	c	0.451	0.942	0.553	1.023	0.564	1.216	0.645	1.269	0.547	1.394
Macapá	k	0.001	9.241	$<$0.000	9.031	0.002	8.895	0.001	8.124	0.177	7.010
	θ	2.172	10.642	3.427	6.153	1.446	7.916	1.719	6.427	0.134	6.737
	c	0.753	1.707	1.023	1.501	0.751	1.719	0.864	1.735	0.492	1.372
Manaus	k	4.197	12.519	3.116	12.289	2.602	11.206	1.959	11.751	1.446	11.773
	θ	0.001	0.088	0.002	0.283	0.003	0.371	0.004	0.715	0.007	1.184
	c	0.311	0.477	0.348	0.591	0.371	0.638	0.382	0.731	0.398	0.831
Palmas	k	1.678	12.520	1.111	11.196	0.765	11.164	1.028	10.752	0.855	9.556
	θ	0.004	1.605	0.010	2.453	0.010	2.895	0.008	1.561	0.027	2.085
	c	0.349	0.758	0.384	0.893	0.381	0.926	0.388	0.832	0.462	1.028
Rio Branco	k	0.139	10.524	0.480	9.173	0.055	8.750	0.163	8.152	0.006	8.718
	θ	0.071	9.786	0.028	3.668	0.172	5.676	0.082	3.974	0.493	3.270
	c	0.432	1.461	0.426	1.064	0.540	1.503	0.504	1.274	0.727	1.448

Assim, examinando os dados do tamanho das sequências de dias chuvosos com limiar 5 mm/dia de chuvas na Tabela 10 o teste de aderência mostrou que as distribuições $GaDTZ(k, \theta)$, $WeDTZ(c, q)$ e $LNDTZ(\mu, \sigma)$ apresentaram-se bem ajustadas ao tamanho das sequências de dias chuvosos para todas as estações, enquanto que as distribuições $GaDTZ(k)$ e $GeoTZ(q)$ não foram adequadas em todas as capitais, isto é, H_0 foi rejeitada para as estações de Belém, Boa Vista, Macapá e Rio Branco para a $GeoTZ(q)$.

Com auxílio dos critérios de informação apresentados na Tabela 10 pode-se constatar que as distribuições $WeDTZ(c, q)$ e $GaDTZ(k, \theta)$ apresentaram os melhores ajustes para maioria das estações, exceto para estação de Manaus em que o critério AIC para distribuição $LNDTZ(\mu, \sigma)$ apresentou o menor valor, mas o critério BIC foi menor para $WeDTZ(c, q)$.

Tabela 10 – Teste qui-quadrado de aderência e critérios de informação para os modelos ajustados no tamanho das sequências de dias chuvosos das capitais da Região Norte de 1961 a 2018 - limiar 5.

Estação	Estatística do teste	Distribuição				
		GaDTZ	GaD1TZ	WeDTZ	LNDTZ	GeoTZ
Belém	χ^2	10.4037	2029.015	7.7854	13.5275	216.1388
	p valor	0.4944	0.0000	0.7323	0.1956	<0.0000
Boa Vista	χ^2	4.3427	59.3442	4.1596	5.8160	28.1182
	p valor	0.6303	<0.0000	0.6550	0.3245	0.0002
Macapá	χ^2	6.6123	193.4153	5.5133	7.5855	28.3937
	p valor	0.5789	<0.0000	0.7015	0.3705	0.0008
Manaus	χ^2	3.5159	11.8767	2.9595	2.3613	9.7188
	p valor	0.6209	0.0647	0.7062	0.6696	0.1370
Palmas	χ^2	4.3710	5.7235	4.0576	1.8501	5.2848
	p valor	0.3581	0.3340	0.3982	0.6040	0.3821
Rio Branco	χ^2	5.2159	8.4442	5.0936	6.1862	15.8885
	p valor	0.2658	0.1333	0.2778	0.1028	0.0071
Estação	Critério de Informação	Distribuição				
		GaDTZ	GaD1TZ	WeDTZ	LNDTZ	GeoTZ
Belém	AIC	11343.19	11709.72	11342.40	11352.87	11469.16
	BIC	11355.82	11716.03	11355.03	11371.81	11475.47
Boa Vista	AIC	3841.414	3873.941	3841.276	3845.544	3858.393
	BIC	3879.339	3852.072	3861.737	3869.264	3863.790
Macapá	AIC	7701.960	7786.367	7701.667	7709.858	7714.433
	BIC	7713.862	7792.318	7713.570	7727.711	7720.385
Manaus	AIC	7215.183	7222.445	7213.583	7211.996	7220.748
	BIC	7227.438	7228.573	7225.838	7230.379	7226.876
Palmas	AIC	2510.863	2509.971	2510.468	2510.078	2509.707
	BIC	2521.017	2515.048	2520.621	2525.308	2514.784
Rio Branco	AIC	4853.467	4853.735	4853.331	4856.975	4860.047
	BIC	4865.211	4859.607	4865.075	4874.592	4865.919

5.1.3 Região Centro-Oeste

As observações das séries históricas de precipitação em relação às estações da Região Centro-Oeste indicaram uma grande variabilidade dos dados segundo o coeficiente de variação, logo a moda e os quartis são as melhores medidas para representar os dados, assim, a moda registrada foi um dia de chuva, isto é, sem dias consecutivos de chuvas, igualmente, a maioria da mediana das estações e limiares estudados. Nota-se que a variabilidade dos dados estão no valor acima do terceiro quartil (Tab. 11).

Tabela 11 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas de 1961 a 2018 das estações meteorológicas das capitais da Região Centro-Oeste.

Capital	Limiar	n	Mo	Q_1	Md	$\overline{x}$	Q_3	Max	SD	CV (%)
Brasília	1	5969	1	1	2	2.32	3	16	1.97	84.95
	2	5248	1	1	1	2.05	3	16	1.65	80.30
	3	4695	1	1	1	1.88	2	14	1.44	76.67
	4	4302	1	1	1	1.76	2	12	1.29	73.36
	5	3937	1	1	1	1.68	2	12	1.19	70.99
Campo Grande	1	3085	1	1	1	1.46	2	7	0.83	57.13
	2	2733	1	1	1	1.40	2	7	0.76	54.17
	3	2482	1	1	1	1.36	2	7	0.72	53.01
	4	2286	1	1	1	1.32	1	7	0.68	51.05
	5	2106	1	1	1	1.29	1	6	0.61	47.56
Cuiabá	1	4624	1	1	1	1.80	2	16	1.39	76.93
	2	4052	1	1	1	1.66	2	16	1.19	71.76
	3	3589	1	1	1	1.56	2	10	1.02	65.21
	4	3258	1	1	1	1.48	2	8	0.89	60.09
	5	2995	1	1	1	1.44	1	7	0.85	59.04
Goiânia	1	6277	1	1	2	2.24	3	18	1.89	84.44
	2	5537	1	1	1	1.99	2	14	1.56	78.62
	3	4997	1	1	1	1.85	2	13	1.36	73.24
	4	4568	1	1	1	1.74	2	13	1.26	72.31
	5	4210	1	1	1	1.65	2	10	1.12	67.94

Legenda: Número de dias chuvosos (n); moda (Mo); $1^{\underline{o}}$ quartil (Q_1); mediana (Md); média ($\overline{x}$); $3^{\underline{o}}$ quartil (Q_3); valor máximo (Max); desvio padrão (sd); coeficiente de variação (cv).

A estação de Campo Grande mostrou ter as menores sequências máximas do tamanho dias chuvosos. Já a estação de Goiânia revelou ter o maior tamanho das sequências de dias chuvosos e variado valores para diferentes limiares (Tab. 11 e Fig. 11).

A validade das suposições de que as séries históricas do tamanho das sequências de dias chuvosos das capitais da Região Centro-Oeste mostraram-se independentes e identicamente distribuídas.

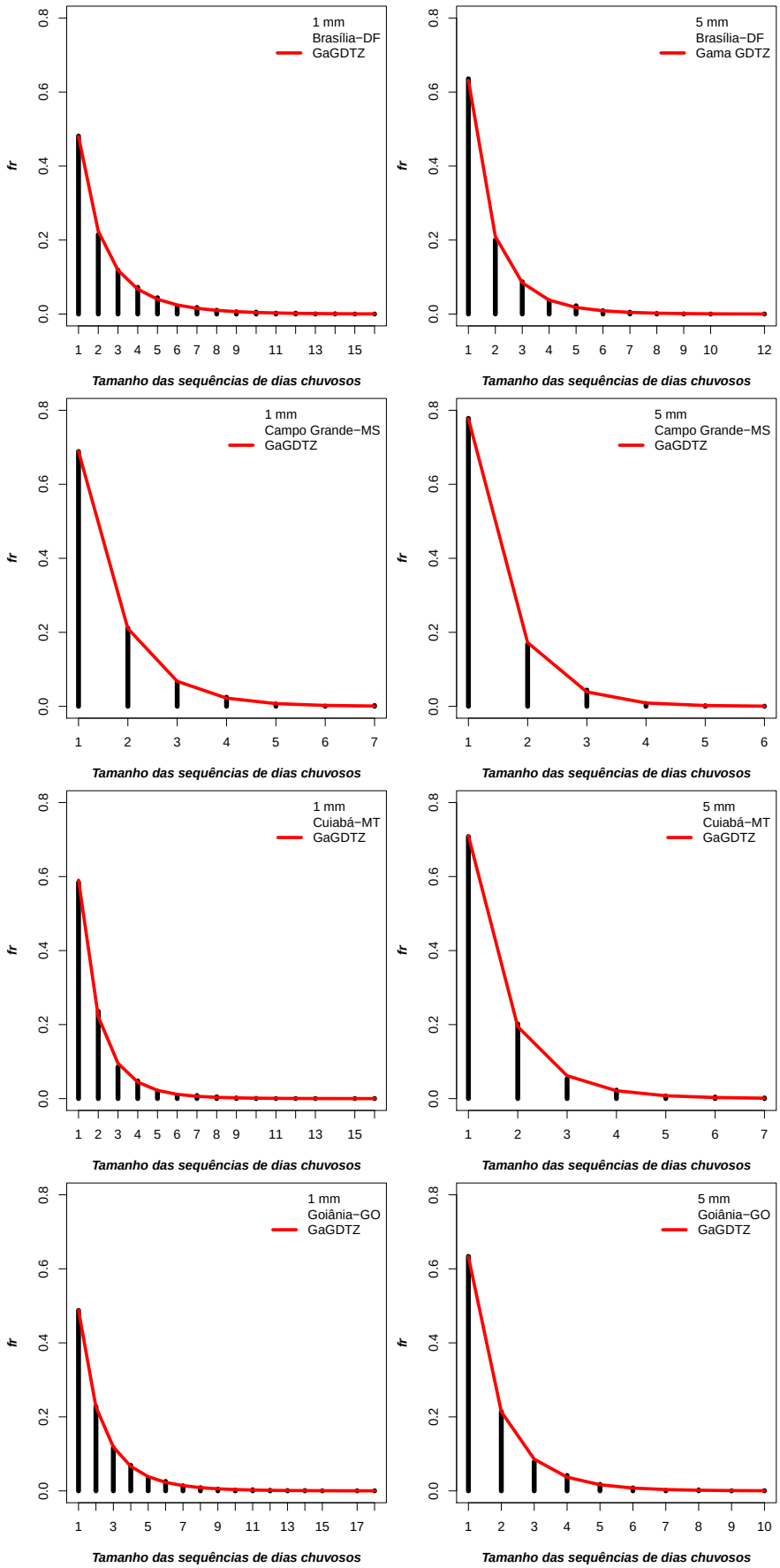

Figura 11 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Centro-Oeste de 1961 a 2018.

A Tabela 12 apresenta os resultados do teste de aderência aplicado nas séries históricas das estações das capitais do Centro-Oeste, podendo notar que a distribuição $GaGDTZ(k, \theta, c)$ não foi rejeitada para nenhuma das estações em diferentes limiares, com nível de significância de 1%, ou seja, o comportamento probabilístico do tamanho das sequências de dias chuvosos pode ser bem ajustados pela distribuição para todos os cenários propostos no estudo, mas ressalta-se que na estação de Cuiabá nos limiares 1 e 2 apresentaram p valores próximos da rejeição.

Observando a Figura 11, pode-se perceber que a distribuição $GaGDTZ(k, \theta, c)$ se moldou aos dados observados independente do limiar utilizado e das estações das capitais da Região Centro-Oeste.

Tabela 12 – Teste qui-quadrado de aderência para a distribuição $GaGDTZ(k, \theta, c)$ ajustada ao tamanho das sequências de dias chuvosos das capitais da Região Centro-Oeste de 1961 a 2018.

Estação	Estatística do teste	Limiar 1	2	3	4	5
Brasília	χ^2	10.1209	10.0722	7.1622	12.0236	6.3148
	p valor	0.4299	0.2599	0.4122	0.0614	0.2768
Campo Grande	χ^2	1.7192	0.8452	0.2665	1.1300	1.6624
	p valor	0.6326	0.6553	0.8752	0.5683	0.1973
Cuiabá	χ^2	16.1129	13.9269	5.9082	3.2455	4.8233
	p valor	0.0241	0.0304	0.2061	0.5176	0.1852
Goiânia	χ^2	5.4098	13.9054	10.1998	10.0972	4.0498
	p valor	0.8621	0.0528	0.0697	0.0725	0.3992

Analisando as estimativas intervalares do parâmetro k, na Tabela 13, pode-se observar que o valor 1 está contido nos intervalos. Contudo a amplitude do intervalo de confiança do parâmetro k foi longa apenas para estação de Cuiabá, isso refletiu na redução da distribuição $GaGDTZ(k, \theta, c)$ na distribuição $LNDTZ(\mu, \sigma)$, observou-se também que as distribuições $GaDTZ(k, \theta)$ e $WeDTZ(q, c)$ apresentaram ajustes adequados ao tamanho das sequências de dias chuvosos para todas as capitais da Região Centro-Oeste (Tab. 14).

Com auxílio dos critérios de informações AIC e BIC para identificar o melhor desempenho do ajuste entre as distribuições, constatou-se que $GaDTZ(k, \theta)$ foi adequada no ajuste do tamanho das sequências de dias chuvosos das estações das capitais, juntamente com valores próximos da distribuição $WeDTZ(q, c)$, exceto para estação Campo Grande que a mais adequada para o ajuste foi a distribuição $GeoTZ(q)$, provavelmente devido apresentar a menor amplitude do tamanho das sequências de dias chuvosos (Tab. 14).

Tabela 13 – Estimativas dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada do tamanho das sequências de dias chuvosos das capitais da Região Centro-Oeste de 1961 a 2018.

Estação	Parâmetros	Limiar 1		Limiar 2		Limiar 3		Limiar 4		Limiar 5	
		LI	LS	LI	LS	LI	LS	LI	LS	LI	LS
Brasília	k	0.609	1.673	0.599	1.702	0.586	1.738	0.572	1.782	0.555	1.835
	θ	1.085	2.954	0.911	2.497	0.795	2.216	0.708	2.040	0.647	1.931
	c	0.838	1.216	0.839	1.215	0.838	1.216	0.834	1.222	0.830	1.229
Campo Grande	k	0.589	1.730	0.591	1.724	0.646	1.577	0.645	1.581	0.598	1.705
	θ	0.673	1.514	0.689	1.479	0.743	1.372	0.749	1.360	0.723	1.411
	c	0.934	1.278	0.981	1.341	1.041	1.361	1.074	1.408	1.095	1.499
Cuiabá	k	2.129	10.125	1.161	10.018	0.530	9.632	0.002	7.978	0.029	5.466
	θ	0.003	0.382	0.006	1.036	0.019	2.461	0.720	2.895	0.312	2.812
	c	0.351	0.606	0.371	0.739	0.421	0.978	0.805	1.440	0.658	1.319
Goiânia	k	0.006	6.710	0.049	3.799	$<$0.000	1.099	$<$0.000	1.781	$<$0.000	1.927
	θ	0.983	8.075	0.482	5.471	2.453	3.417	1.748	2.762	1.620	2.592
	c	0.670	1.568	0.603	1.329	1.059	1.421	0.916	1.263	0.939	1.314

Tabela 14 – Teste qui-quadrado de aderência e critérios de informação para os modelos ajustados da série histórica das estação das capitais da Região Centro-Oeste de 1961 a 2018 - limiar 5.

Estação	Estatística do teste	Distribuição				
		GaDTZ	GaD1TZ	WeDTZ	LNDTZ	GeoTZ
Brasília	χ^2	5.5311	66.2620	7.1142	12.8033	44.5768
	p valor	0.3545	<0.0000	0.2123	0.0122	<0.0000
Campo Grande	χ^2	1.4799	7.4395	1.5332	3.3177	1.6402
	p valor	0.4771	0.0591	0.4646	0.1904	0.6503
Cuiabá	χ^2	5.1787	8.0896	4.6784	3.9514	17.4554
	p valor	0.2694	0.1513	0.3219	0.2667	0.0037
Goiânia	χ^2	4.3562	34.0024	5.3773	10.9545	24.5491
	p valor	0.4993	<0.0000	0.3715	0.0270	0.0004
Estação	Critério de Informação	Distribuição				
		GaDTZ	GaD1TZ	WeDTZ	LNDTZ	GeoTZ
Brasília	AIC	5279.463	5328.994	5281.441	5290.707	5314.929
	BIC	5290.983	5334.754	5292.961	5307.987	5320.689
Campo Grande	AIC	2247.452	2253.335	2247.503	2249.644	2245.582
	BIC	2258.248	2258.733	2258.299	2260.440	2250.980
Cuiabá	AIC	3681.849	3681.605	3681.642	3684.046	3688.607
	BIC	3693.129	3687.245	3692.922	3700.966	3694.247
Goiânia	AIC	5625.483	5651.583	5626.492	5634.688	5644.407
	BIC	5637.175	5657.429	5638.183	5652.225	5650.253

5.1.4 Região Sudeste

Para as estações das capitais da região Sudeste, a análise exploratória indicou que os coeficientes de variações foram maiores que 50%, ou seja, sinalizando que a moda e os quartis são as melhores medidas para representar os dados. Logo, salienta-se que a mediana mostrou que 50% dos dias chuvosos, das séries históricas, não tiveram dias consecutivos de chuva, assim como também a moda (Tab. 15).

Tabela 15 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas de 1961 a 2018 das estações meteorológicas das capitais da região Sudeste.

Capital	Limiar	n	Mo	Q_1	Md	$\overline{x}$	Q_3	Max	SD	CV (%)
Belo Horizonte	1	4789	1	1	1	2.23	3	16	1.93	86.80
	2	4312	1	1	1	2.11	3	15	1.81	85.41
	3	3965	1	1	1	2.02	2	15	1.68	83.01
	4	3676	1	1	1	1.91	2	15	1.53	79.88
	5	3413	1	1	1	1.82	2	15	1.42	77.73
Rio de Janeiro	1	3101	1	1	1	1.90	2	13	1.28	67.51
	2	2665	1	1	1	1.75	2	8	1.12	63.84
	3	2302	1	1	1	1.64	2	8	1.01	61.68
	4	2017	1	1	1	1.53	2	8	0.87	56.80
	5	1795	1	1	1	1.48	2	8	0.82	55.25
São Paulo	1	5956	1	1	1	1.99	2	15	1.50	74.98
	2	5252	1	1	1	1.82	2	15	1.28	70.64
	3	4725	1	1	1	1.71	2	15	1.15	67.24
	4	4349	1	1	1	1.64	2	10	1.06	64.89
	5	3984	1	1	1	1.58	2	10	0.99	63.30
Vitória	1	4622	1	1	1	2.06	3	13	1.55	75.49
	2	3893	1	1	1	2.37	2	13	1.36	72.44
	3	3367	1	1	1	1.76	2	10	1.21	68.92
	4	2930	1	1	1	1.64	2	10	1.09	66.29
	5	2628	1	1	1	1.56	2	10	0.99	63.78

Legenda: Número de dias chuvosos (n); moda (Mo); $1^{\underline{o}}$ quartil (Q_1); mediana (Md); média ($\overline{x}$); $3^{\underline{o}}$ quartil (Q_3); valor máximo (Max); desvio padrão (sd); coeficiente de variação (cv).

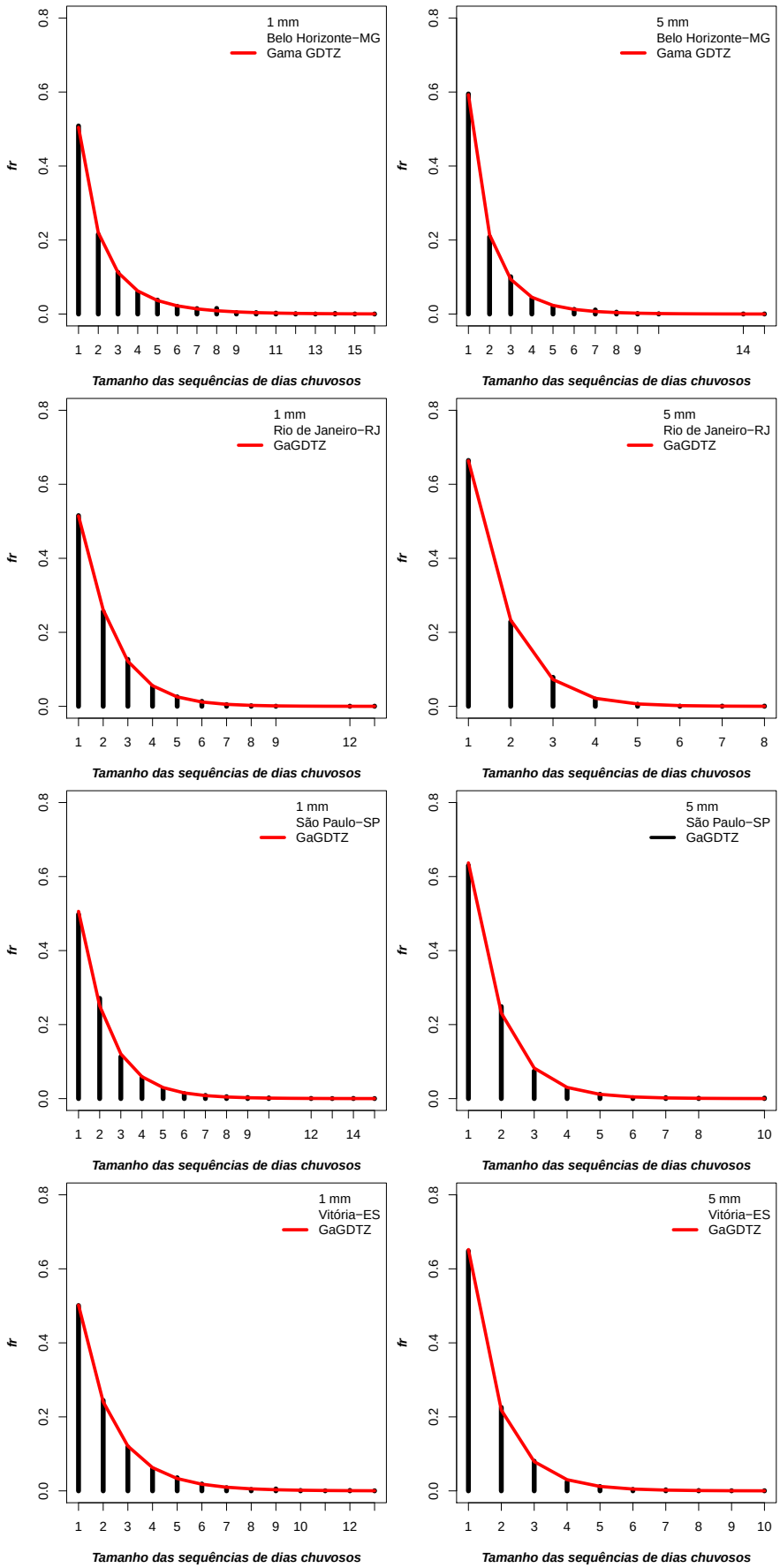

Figura 12 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Sudeste de 1961 a 2018.

Quanto ao tamanho máximo das sequências de dias chuvosos, com limiar 1 mm/dia, observou-se que a estação do Rio de Janeiro apresentou as maiores sequências dias chuvosos entre as capitais, isto é, com 21 dias de chuvas consecutivos. No entanto, ao aumentar o limiar para 5 mm/dia o tamanho máximo das sequências de dias com chuvas reduziu para 9 dias consecutivos, ressalta-se que esse total sucessivo de dias de chuvas pode ser catastrófico para comunidade do Rio de Janeiro (Tab. 15 e Fig. 12).

Notou-se que as estações de Belo Horizonte e São Paulo não apresentaram variações no tamanho máximo de dias chuvosos com a mudança dos limiares 2, 3, 4 ou 5 mm/dia, também, na estação de Vitória não teve variação na sequência máxima para os limiares 3, 4 ou 5 mm/dia. Todos esses resultados podem ser um indicativo que essas sequências máximas de chuvas não foram compostas por chuvas fracas (Tab. 15 e Fig. 12).

Para realizar as estimativas de máxima verossimilhança foram analisadas as séries históricas do tamanho das sequências de dias chuvosos das capitais da Região Sudeste por meio dos testes Wald-Wolfowitz e Mann-Kendall, com nível de 1% de significância, as observações das séries mostraram-se independentes e identicamente distribuídas.

Avaliando o teste χ^2 de aderência empregado nas séries históricas das estações das capitais do Sudeste (Tab. 16), pode-se constatar que a distribuição $GaDGTZ(k, \theta, c)$, com nível de significância de 1%, não foi rejeitada para nenhuma das estações e limiares, mas na estação de São Paulo nos limiares 4 e 5 apresentaram p valores pequenos. Tal resultado pode ser visualizado na Figura 12, isto é, pode-se notar que a distribuição $GaDGTZ(k, \theta, c)$ aparentemente se adequou ao tamanho das sequências de chuvas das estações de todas as capitais do Sudeste, independente do limiar utilizado.

Tabela 16 – Teste qui-quadrado de aderência para a distribuição $GaGDTZ(k, \theta, c)$ ajustada ao tamanho das sequências de dias chuvosos das capitais da Região Sudeste de 1961 a 2018.

Capital	Estatística do teste	Limiar 1	2	3	4	5
Belo Horizonte	χ^2	2.3483	2.4645	6.0017	5.6784	3.2090
	p valor	0.5033	0.4817	0.1115	0.0584	0.2009
Rio de Janeiro	χ^2	1.6020	2.3519	0.9579	1.5855	1.0203
	p valor	0.8084	0.6713	0.8114	0.4525	0.6004
São Paulo	χ^2	10.7779	9.6127	4.9155	9.4317	6.8017
	p valor	0.1486	0.1419	0.4262	0.0511	0.0333
Vitória	χ^2	6.1232	6.1042	5.6290	5.5560	3.9097
	p valor	0.5254	0.4116	0.3439	0.2348	0.4183

Tabela 17 – Estimativas dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada do tamanho das sequências de dias chuvosos das capitais da Região Sudeste de 1961 a 2018.

Estação	Parâmetros	Limiar 1		Limiar 2		Limiar 3		Limiar 4		Limiar 5	
		LI	LS	LI	LS	LI	LS	LI	LS	LI	LS
Belo Horizonte	k	<0.000	8.169	<0.000	7.766	<0.000	7.461	<0.000	6.654	<0.000	7.875
	θ	2.635	4.119	2.134	3.617	1.829	3.311	1.728	3.091	1.419	2.764
	c	0.874	1.215	0.819	1.162	0.791	1.148	0.814	1.181	0.777	1.150
Rio de Janeiro	k	0.554	10.504	0.237	7.624	0.406	7.307	0.189	7.203	0.141	6.678
	θ	0.050	5.427	0.162	5.700	0.118	3.701	0.235	4.536	0.276	4.197
	c	0.467	1.359	0.559	1.638	0.546	1.400	0.657	1.834	0.686	1.822
São Paulo	k	3.777	10.086	3.160	8.096	1.835	8.166	2.263	8.398	3.141	8.661
	θ	0.008	0.209	0.021	0.279	0.028	0.810	0.022	0.498	0.017	0.241
	c	0.411	0.607	0.468	0.667	0.482	0.828	0.475	0.759	0.474	0.682
Vitória	k	0.084	9.990	0.059	9.098	0.022	8.619	0.329	8.815	0.475	8.942
	θ	0.157	11.790	0.204	8.790	0.396	6.999	0.047	3.955	0.037	2.862
	c	0.471	1.761	0.516	1.693	0.610	1.795	0.462	1.180	0.467	1.106

Considerando, as estimativas intervalares dos parâmetros com limiar 5, notou-se em comum que os intervalos de confiança do parâmetro c contém o valor 1 para a maioria das estações com exceção da estação de São Paulo (Tab. 17). Observando a Tabela 18, pode-se constatar que a distribuição $GaDTZ(k, \theta)$ e $WeDTZ(q, c)$, apresentaram um bom desempenho no ajuste do tamanho das sequências de chuvas das estações de Belo Horizonte, Rio de Janeiro e Vitória. Contudo, a estação de São Paulo não seguiu o mesmo resultado, pois apenas a distribuição $LNDTZ(\mu, \sigma)$ não foi rejeitada, com nível de 1% de significância, no teste de aderência, como já sinalizado na estimação intervalar.

Para discriminar a melhor distribuição dos casos particulares da distribuição $GaGDTZ$ (k, θ, c), os critérios de informações AIC e BIC apresentaram valores muito próximos entre as distribuições, indicando que $GaDTZ(k, \theta)$ e $WeDTZ(q, c)$ forneceram os melhores ajustes no tamanho das sequências de dias chuvosos para todas as capitais da Região Sudeste, exceto em São Paulo que a distribuição $LNDTZ(\mu, \sigma)$ apresentou melhor desempenho (Tab. 18).

Tabela 18 – Teste qui-quadrado de aderência e critério de informação para os modelos ajustados no tamanho das sequências de dias chuvosos das capitais da Região Sudeste de 1961 a 2018 - limiar 5.

Estação	Estatística do teste	Distribuição				
		GaDTZ	GaD1TZ	WeDTZ	LNDTZ	GeoTZ
Belo Horizonte	χ^2	10.7521	135.8446	10.0219	13.5897	56.2021
	p-value	0.0963	<0.0000	0.1237	0.0345	<0.0000
Rio de Janeiro	χ^2	0.8620	5.5951	0.8180	2.7913	3.0777
	p-value	0.8345	0.2314	0.8451	0.2476	0.5448
São Paulo	χ^2	20.0311	22.0399	18.4253	6.4213	22.0624
	p-value	0.0012	0.0011	0.0024	0.2674	0.0011
Vitória	χ^2	4.9842	9.9027	4.1658	1.8039	10.7723
	p-value	0.4178	0.1288	0.5257	0.7717	0.0956
Estação	**Critério de Informação**	**Distribuição**				
		GaDTZ	GaD1TZ	WeDTZ	LNDTZ	GeoTZ
Belo Horizonte	AIC	4662.795	4744.184	4661.948	4668.088	4701.659
	BIC	4673.864	4749.718	4673.016	4677.157	4707.193
Rio de Janeiro	AIC	2268.309	2270.800	2268.419	2271.414	2268.155
	BIC	2278.506	2275.898	2278.615	2286.709	2273.253
São Paulo	AIC	5237.684	5236.154	5236.718	5225.103	5236.262
	BIC	5249.352	5241.988	5248.386	5236.771	5242.096
Vitória	AIC	3436.961	3437.715	3436.285	3436.408	3438.569
	BIC	3447.814	3443.142	3447.139	3452.688	3443.996

5.1.5 Região Sul

As séries históricas das capitais da Região Sul mostraram que a moda do tamanho das sequências de dias chuvosos foi de um dia, independente do limiar utilizado e estação estudada. Observou-se também que o maior tamanho das sequências de dias chuvosos no ano foi de 21 dias, ocorrido em Florianópolis, com limiar 1 mm/dia de chuvas. Ao mudar de limiar, Florianópolis foi a estação que apresentou mais diferenças entre o tamanho máximo das sequências de dias chuvosos, enquanto que, na estação de Curitiba, o tamanho máximo das sequências de dias chuvosos foi de 10 dias chuvosos para a maioria dos limiares utilizados, assim como Porto Alegre (Tabela 19).

No estudo de Li et al. (2017) caracterizando os períodos úmidos e secos da China, tais pesquisadores declararam que o período de 1 dia de chuva ou seca dominava sobre todas as zonas climáticas do país. Nas quatro estações meteorológicas da região do delta de Mahanadi na Índia (Bhubaneswar, Cuttack, Paradip e Puri) avaliadas por Sukla et al. (2012), também apresentaram variabilidade alta e moda de um dia chuvoso. O mesmo foi constatado por Deni et al. (2010), Deni e Jemain (2009), Deni et al. (2008) em que as estações pluviométricas avaliadas apresentaram moda 1 e variabilidade alta no tamanho de dias chuvosos. No trabalho de Zolina et al. (2013) estações europeias também apresentaram moda 1 dia chuvoso. Nesse trabalho, também, em todas as estações o tamanho das sequências de dias chuvosos teve moda 1 e variabilidade grande.

Tabela 19 – Análise descritiva do tamanho das sequências de dias chuvosos da série histórica de 1961 a 2018 das estações meteorológicas das capitais da região Sul.

Capital	Limiar	n	Mo	Q_1	Md	$\overline{x}$	Q_3	Max	sd	cv (%)
Curitiba	1	6062	1	1	1	1.94	2	15	1.36	70.02
	2	5219	1	1	1	1.78	2	10	1.18	66.13
	3	4594	1	1	1	1.67	2	10	1.07	63.89
	4	4181	1	1	1	1.60	2	10	0.98	61.29
	5	3845	1	1	1	1.54	2	10	0.91	59.01
Florianópolis	1	5257	1	1	1	1.98	2	21	1.49	75.60
	2	4507	1	1	1	1.80	2	12	1.21	67.53
	3	3993	1	1	1	1.70	2	11	1.10	65.08
	4	3601	1	1	1	1.63	2	11	1.03	63.39
	5	3268	1	1	1	1.57	2	9	0.94	59.83
Porto Alegre	1	5907	1	1	1	1.80	2	10	1.15	63.74
	2	5152	1	1	1	1.67	2	10	1.01	60.47
	3	4617	1	1	1	1.58	2	10	0.92	58.45
	4	4166	1	1	1	1.52	2	9	0.86	56.71
	5	3835	1	1	1	1.48	2	9	0.82	55.71

Legenda: Número de dias chuvosos (n); moda (Mo); $1^{\underline{o}}$ quartil (Q_1); mediana (Md); média ($\overline{x}$); $3^{\underline{o}}$ quartil (Q_3); valor máximo (Max); desvio padrão (sd); coeficiente de variação (cv).

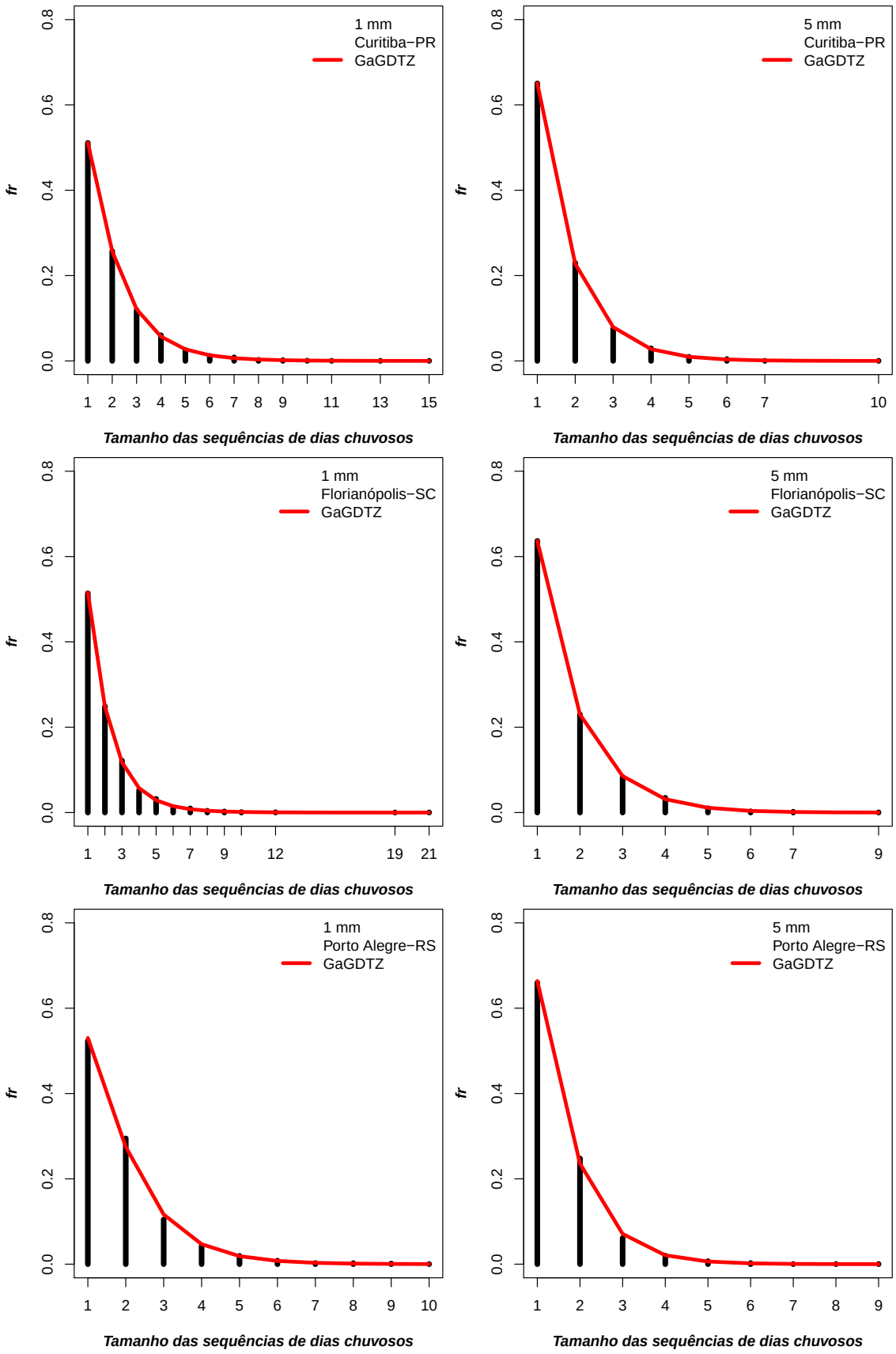

Figura 13 – Ajuste da distribuição $GaGDTZ(k, \theta, c)$ ao tamanho das sequências de dias chuvosos das séries históricas das estações das capitais da Região Sul de 1961 a 2018.

Para todas as estações, a validade das suposições mostraram-se independentes e identicamente distribuídas para as observações das séries históricas do tamanho das sequências de dias chuvosos das capitais da região Sul, ou seja, os testes Wald-Wolfowitz e Mann-Kendall não rejeitaram a hipótese H_0.

As distribuições observadas e ajustadas do tamanho das sequências de chuvas das estações da Região Sul estão apresentadas na Figura 13. Pode-se notar que a distribuição $GaGDTZ(k, \theta, c)$, independente do limiar e estação, aparentemente, mostrou-se adequadamente ajustada aos dados.

O teste χ^2 de aderência corrobora para justificar essas observações ilustrados nos gráficos de ajustes, pois independente da estação e limiar utilizado a distribuição $GaGDTZ(k, \theta, c)$ foi adequada para ajustar os dados, com nível de significância 1% (Tabela 20).

Tabela 20 – Teste qui-quadrado de aderência para a distribuição $GaGDTZ(k, \theta, c)$ ajustada as séries históricas das estações meteorológicas das capitais da região Sul de 1961 a 2018.

Estação	Estatística do teste	Limiar				
		1	2	3	4	5
Curitiba	χ^2	3.5197	4.2973	1.2774	3.8197	2.1703
	p valor	0.7413	0.5074	0.8651	0.4309	0.1407
Florianópolis	χ^2	6.0028	1.7566	2.1914	1.4546	3.1586
	p valor	0.5394	0.9406	0.8221	0.8346	0.3678
Porto Alegre	χ^2	14.5127	7.8197	9.0548	5.8574	5.5008
	p valor	0.0185	0.0984	0.0286	0.1188	0.1385

A Tabela 21 apresenta as estimativas intervalares dos parâmetros para os diferentes limiares, mas para realizar as comparações das estimativas dos parâmetros e dos casos particulares da distribuição $GaGDTZ(k, \theta, c)$, tomou-se os intervalos de confiança do limiar 5 mm/dia de chuva e foi constatado que o parâmetro c contém o valor 1 nas estações de Curitiba e Florianópolis, enquanto que na estação de Porto Alegre não contém, mas nota-se uma amplitude longa da estimativa intervalar do parâmetro k, podendo constatar que isso reflete nos ajustes adequados indicados pelo teste de aderência, ou seja, houve bons ajustes com distribuições $GaDTZ(k, \theta)$, $GeoTZ(q)$, $GaDTZ(k)$, assim como para a distribuição $WeDTZ(q, c)$ que nas estimativas intervalares do parâmetro k contém valor 1, bem como a distribuição $(LNDTZ(\mu, \sigma))$ que a amplitude intervalar desse parâmetro mostrou-se ser longa (Tab. 22).

Tabela 21 – Estimativas dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada do tamanho das sequências de dias chuvosos das capitais da Região Sul de 1961 a 2018.

Estação	Parâmetros	Limiar 1		Limiar 2		Limiar 3		Limiar 4		Limiar 5	
		LI	LS	LI	LS	LI	LS	LI	LS	LI	LS
Curitiba	k	1.006	10.238	0.459	9.133	0.096	9.000	0.239	8.773	0.106	8.415
	θ	0.026	2.983	0.063	4.702	0.184	6.826	0.104	4.954	0.189	5.289
	c	0.441	1.050	0.488	1.332	0.555	1.786	0.534	1.523	0.593	1.751
Florianópolis	k	1.156	13.078	1.004	9.897	0.148	9.133	0.306	8.659	0.081	8.610
	θ	0.006	2.804	0.026	2.464	0.132	6.643	0.080	4.523	0.227	5.923
	c	0.365	0.895	0.449	1.027	0.525	1.653	0.509	1.382	0.599	1.880
Porto Alegre	k	4.843	11.878	3.965	10.511	3.055	9.778	2.885	9.211	2.712	9.766
	θ	0.011	0.196	0.021	0.298	0.030	0.481	0.031	0.450	0.025	0.477
	c	0.469	0.680	0.511	0.754	0.540	0.849	0.542	0.837	0.526	0.844

Utilizando os critérios de informações AIC e BIC para verificar o melhor desempenho da distribuição, notou-se que para as estações de Curitiba e Florianópolis a distribuição $GaDTZ(k)$ e a $GeoTZ(q)$ foram mais adequadas no ajuste do tamanho das sequências de dias chuvosos das estações, enquanto que para a estação de Porto Alegre foi a distribuição $LNDTZ(\mu, \sigma)$ (Tab. 22).

É importante ressaltar que as estações com amplitudes do tamanho das sequências superiores a 12 dias sucessivos de chuvas não tiveram ajustes adequados para as distribuições $GaDTZ(k)$ e a $GeoTZ(q)$, mas pode-se constatar que tais distribuições apresentaram melhores ajustes em estações com amplitudes do tamanho das sequências inferiores a 13 dias sucessivos de chuvas salvo, por exemplo as estações de Porto Alegre, São Paulo e Teresina, em razão das frequências dos valores observados diferirem dos valores esperados dessas duas distribuições, logo a adequação dos ajustes das distribuições $GaDTZ(k)$ e $GeoTZ(q)$ dependem da amplitude curta do tamanho das sequências dias chuvosos e dos valores esperados próximos dos valores observados.

No trabalho de Deni et al. (2008) o ajuste da distribuição geométrica não foram adequados para o tamanho das sequências de dias chuvosos nas 10 estações pluviométricas da Península da Malásia avaliadas, ressaltando que a amplitude do tamanho das sequências foram acima de 17 dias sucessivos de chuvas com moda centrado em 1 para todas as estações. Enquanto que as estações pluviométricas avaliadas por Zolina et al. (2013) apresentaram ajustes adequados com a distribuição GeoTZ(q) e a amplitude do tamanho das sequências de dias chuvosos das estações foram inferiores a 15 dias e moda 1.

Tabela 22 – Teste qui-quadrado de aderência e critérios de informação para os modelos ajustados das séries históricas das estação das capitais da região Sul de 1961 a 2018 - limiar 5.

Estação	Estatística do teste	Distribuição				
		GaDTZ	GaD1TZ	WeDTZ	LNDTZ	GeoTZ
Curitiba	χ^2	3.0412	3.3042	3.0380	6.2159	3.0370
	p valor	0.5509	0.6531	0.5514	0.1015	0.6942
Florianópolis	χ^2	2.0440	2.0592	2.0371	6.1622	2.0379
	p valor	0.7276	0.8408	0.7289	0.1039	0.8438
Porto Alegre	χ^2	10.0257	21.7784	11.3754	3.0602	15.2552
	p valor	0.0399	0.0005	0.0226	0.3824	0.0093
Estação	**Critério de Informação**	**Distribuição**				
		GaDTZ	GaD1TZ	WeDTZ	LNDTZ	GeoTZ
Curitiba	AIC	4976.955	4975.340	4976.955	4982.287	4974.958
	BIC	4988.606	4981.166	4988.606	4999.764	4980.784
Florianópolis	AIC	4286.666	4284.762	4286.639	4293.747	4284.700
	BIC	4297.947	4290.402	4297.920	4310.669	4290.341
Porto Alegre	AIC	4832.433	4841.342	4833.913	4826.630	4835.078
	BIC	4844.152	4847.202	4845.632	4844.209	4844.938

5.2 Análises das Séries Mensais

Para o estudo das análises das séries históricas mensais foi utilizado as distribuições dos casos especiais da distribuição $GaGDTZ(k, \theta, c)$ com o melhor ajuste utilizando o princípio da parcimônia, uma vez que as análises exploratórias do tamanho das sequências de dias chuvosos das séries históricas anuais das estações do Porto São José e Porto Rico, com limiar 5 mm/dia, indicaram modas iguais a 1, variabilidades dos dados com coeficientes de variação acima de 40% e amplitudes do tamanho das sequências inferiores a 13 dias (Tab. 23), além do que o teste χ^2 de aderência indicou que a distribuição $GaGDTZ(k, \theta, c)$ apresenta um ajuste adequado ao tamanho das sequências de dias chuvosos das séries históricas anuais dessas duas estações do Rio Paraná (Tab. 24).

Tabela 23 – Análise descritiva do tamanho das sequências de dias chuvosos séries históricas anuais de duas estações do Rio Paraná - limiar 5.

Estação	n	Mo	Q_1	Md	$\overline{x}$	Q_3	Max	sd	cv (%)
Porto São José	1730	1	1	1	1.42	2	7	0.80	56.45
Porto Rico	2552	1	1	1	1.21	1	8	0.50	41.60

Legenda: Número de dias chuvosos (n); moda (Mo); $1^{\underline{o}}$ quartil (Q_1); mediana (Md); média ($\overline{x}$); $3^{\underline{o}}$ quartil (Q_3); valor máximo (Max); desvio padrão (sd); coeficiente de variação (cv).

Ressalta-se que o pressuposto do tamanho das sequências de dias chuvosos independentes e identicamente distribuídos não foram violados, segundo os teste Wald Wolfowitz e Mann-Kendall, com nível de significância de 1%.

Tabela 24 – Teste de aderência e estimativas dos parâmetros da distribuição $GaGDTZ(k, \theta, c)$ ajustada ao tamanho das sequências de dias chuvosos das séries históricas anuais de duas estações do Rio Paraná - limiar 5.

Estação	Estatística do teste	Parâmetro	IC (95%)	
			Inferior	Superior
Porto São José	$\chi^2 = 2.9768$	k	0.065	9.325
	p valor $= 0.2257$	θ	0.130	4.242
	gl $= 2$	c	0.545	1.513
Porto Rico	$\chi^2 = 1.4011$	k	0.005	11.228
	p valor $= 0.2365$	θ	0.389	2.122
	gl $= 1$	c	0.801	1.620

Na Tabela 24 constatou-se também, que as estimativas intervalares dos parâmetros de c, θ, k contém o valor 1, sugerindo que a distribuição $GaGDTZ(k, \theta, c)$ reduz a distribuição $GeoTZ(q)$, assim como, as distribuições: $WeDTZ(q, c)$, $GaDTZ(k, \theta)$ e $GaDTZ(k)$ (Tabela 24). Diante disso, usando o critério de parcimônia e fato das distribuições $GaDTZ(k)$ e $GeoTZ(q)$ serem mais adequadas em séries com amplitude curta do tamanho das sequências de dias chuvosos, deste modo optou-se em utilizá-las para avaliar a adequação dessas distribuições

de probabilidade em ajustar o tamanho das sequências de dias chuvosos mensais, assim como também a distribuição Lindley discreta truncada em zero, por ter apenas um parâmetro.

5.2.1 Estação Porto São José

A análise exploratória das séries históricas mensais do Porto São José indicou que a moda foi 1 para todos os meses, inclusive o terceiro quartil foi 1 nos meses de março, maio, julho, outubro e novembro, indicando que 75% dos dias chuvosos não tiveram sequências de chuvas no dia posterior, enquanto que para os meses de janeiro, fevereiro, abril, junho, agosto, setembro e dezembro a mediana foi de 1 dia chuvoso e o terceiro quartil foi 2 dias chuvosos, evidenciando que 25% dos dias chuvosos tiveram mais de 2 dias consecutivos de chuvas (Tab. 25).

Tabela 25 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas mensais da estação do Porto São José do rio Paraná - 1985 a 2018.

Mês	n	Mo	Q_1	Md	$\overline{x}$	Q_3	Max	SD	CV (%)
Jan	223	1	1	1	1.58	2	6	0.94	59.10
Fev	191	1	1	1	1.53	2	7	1.01	65.75
Mar	121	1	1	1	1.29	1	4	0.68	52.97
Abr	117	1	1	1	1.35	2	3	0.55	40.61
Mai	124	1	1	1	1.23	1	5	0.62	50.05
Jun	100	1	1	1	1.41	2	5	0.79	55.75
Jul	95	1	1	1	1.32	1	4	0.67	50.60
Ago	81	1	1	1	1.56	2	4	0.96	61.52
Set	143	1	1	1	1.43	2	4	0.76	52.83
Out	179	1	1	1	1.37	1	5	0.79	57.56
Nov	167	1	1	1	1.37	1	5	0.76	55.72
Dez	189	1	1	1	1.40	2	5	0.74	52.50

Legenda: Número de dias chuvosos (n); moda (Mo); $1^{\underline{o}}$ quartil (Q_1); mediana (Md); média ($\overline{x}$); $3^{\underline{o}}$ quartil (Q_3); valor máximo (Max); desvio padrão (sd); coeficiente de variação (cv).

Ressalta-se ainda a amplitude curta dos tamanhos das sequências de dias chuvosos, pois o mês de fevereiro apresentou o maior tamanho máximo das sequências de dias chuvosos, isto é, 7 dias consecutivos de chuva, seguido pelo mês de janeiro com o tamanho máximo de 6 dias consecutivos de chuvas e com o maior número de dias chuvosos da série (223). Enquanto que abril foi o mês que apresentou o menor tamanho máximo das sequências de dias chuvosos, ou seja, 3 dias de chuvas consecutivos (Tab. 25 e Figs. 14 e 15).

Para verificar a validade das suposições de independentes e identicamente distribuídas para o tamanho das sequências de dias chuvosos da estação os testes Wald-Wolfowitz e Mann-Kendall foram realizados e indicaram que tais suposições não foram violadas, com nível de significância de 1%.

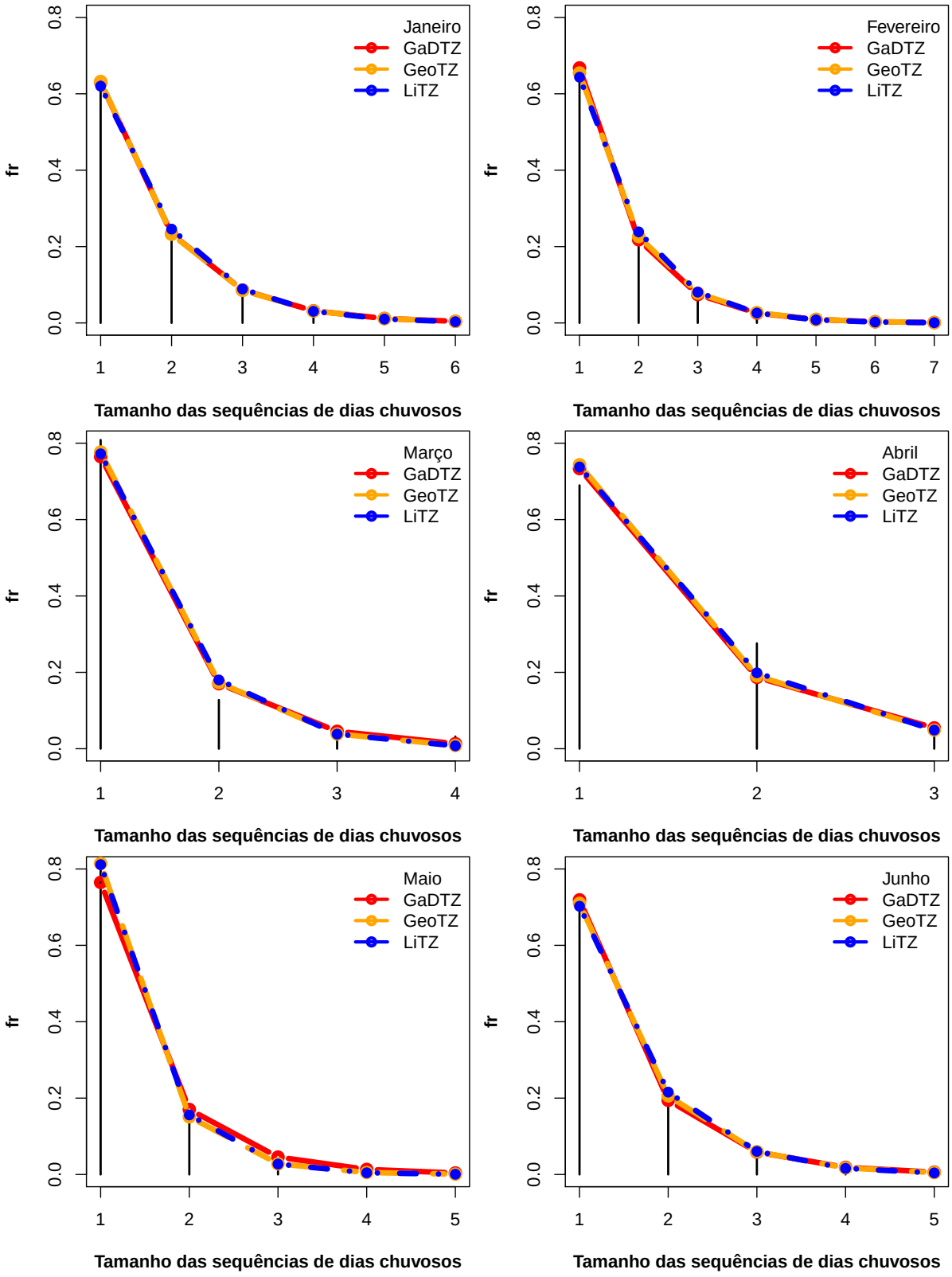

Figura 14 – Ajuste do tamanho das sequências de dias chuvosos das séries de janeiro a junho da estação Porto São José do rio Paraná - 1985 a 2018.

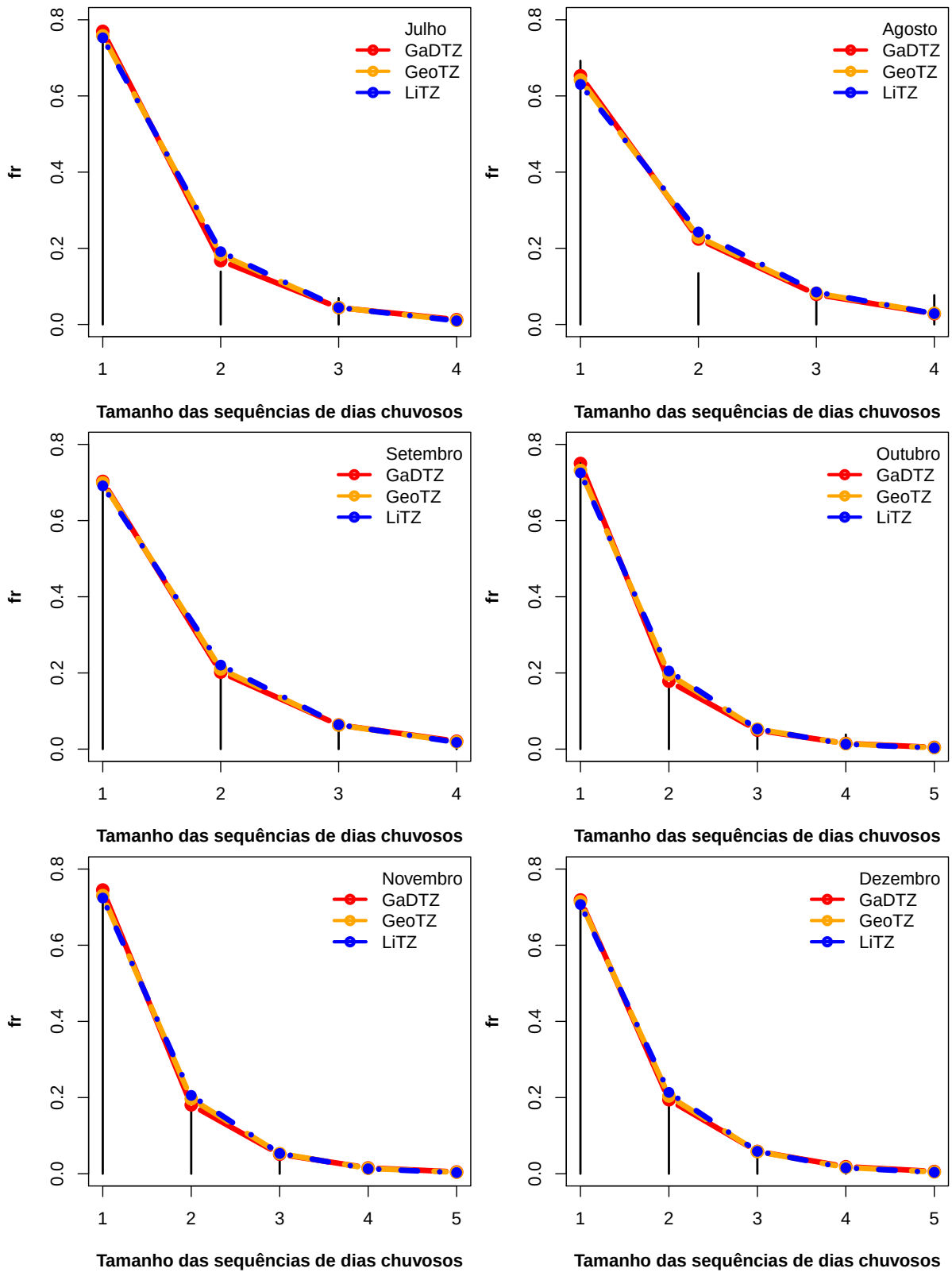

Figura 15 – Ajuste do tamanho das sequências de dias chuvosos das séries de julho a dezembro da estação porto São José do rio Paraná - 1985 a 2018.

O teste χ^2 de aderência mostrou que as distribuições Lindley discreta truncada em zero $(LiDTZ(\beta))$, $GaDTZ(k)$ e a $GeoTZ(q)$ foram adequadas ao ajuste do tamanho das sequências de dias chuvosos mensais, com nível de significância de 1% (Tabela 26).

Tabela 26 – Teste qui-quadrado de aderência para as distribuições ajustadas as séries históricas mensais da estação Porto São José do rio Paraná - 1985 a 2018.

Mês	Estatística do teste	Distribuição		
		LiDTZ	GeTZ	GaDTZ
Jan	χ^2	0.3028	0.0931	0.0861
	p valor	0.8594	0.9545	0.9578
Fev	χ^2	1.0596	0.5128	0.2675
	p valor	0.5887	0.7738	0.8747
Mar	χ^2	6.6232	5.0262	2.4065
	p valor	0.0364	0.0810	0.3002
Abr	χ^2	4.5136	5.4884	6.7239
	p valor	0.1046	0.0642	0.0346
Mai	χ^2	0.2761	0.2052	3.1109
	p valor	0.5992	0.6505	0.0777
Jun	χ^2	0.3039	0.0882	0.0102
	p valor	0.8590	0.9568	0.9948
Jul	χ^2	2.1382	1.8332	1.4715
	p valor	0.3433	0.3998	0.4791
Ago	χ^2	7.8763	7.2084	7.4121
	p valor	0.0486	0.0655	0.0598
Set	χ^2	0.3988	0.1518	0.0759
	p valor	0.8192	0.9269	0.9627
Out	χ^2	2.0085	1.2014	0.3102
	p valor	0.3663	0.5484	0.8562
Nov	χ^2	2.9739	1.8885	0.7532
	p valor	0.2260	0.3889	0.6861
Dez	χ^2	0.0469	0.0183	0.2918
	p valor	0.9767	0.9908	0.8642

É importante ressaltar que as curvas ajustadas nos meses de março, abril e agosto não seguiram os perfis das distribuições observadas, devido as frequências dos valores esperados divergirem das frequências dos valores observados (Figs. 14 e 15). A Tabela 27 de valores observados e esperados ajuda evidenciar melhor tais ocorrências, por exemplo, no mês de março o tamanho 1 foi de 76 dias chuvosos, em que o valor esperado mais próximo do empírico foi a frequência de 73.02 para distribuição $GeoTZ(q)$, já, no tamanho 2 o valor observado foi de 12 dias, mas o esperado mais próximo foi 16.01 para distribuição $GaDTZ(k)$. Observe o mês de junho ocorreu os valores esperados mais próximos dos observados, isto é, o tamanho 1 foi de 51 dias chuvosos e o valor esperado mais próximo foi de 51.06 para distribuição $GaDTZ(k)$, o tamanho 2 do valor observado foi de 14 enquanto o esperado mais próximo foi de 13.79 com mesma distribuição, assim como para tamanho 3 e maior.

Destaca-se também que o p valor mais alto entre os meses foi em junho com a distribuição $GaDTZ(k)$ indicando que quanto maior p valor mais próximos serão os valores esperados dos observados.

Salienta-se ainda que a qualidade do ajuste de cada uma das distribuições foi avaliada pelos critérios de informação derivados do ajuste pelo método da máxima verossimilhança, mostrando que os valores dos critérios AIC e BIC para as três distribuições foram muitos próximos nos meses de janeiro, junho, julho, agosto, setembro e dezembro. Os critérios também indicaram que as distribuições $GaDTZ(k)$ e $GeoTZ(q)$ foram selecionadas mais vezes, uma vez que a distribuição $GaDTZ(k)$, só não apresentou menores valores dos critérios nos meses de abril e maio, já a $GeoTZ(q)$ não foi melhor nos meses de abril e outubro, enquanto que a $LiDTZ(\beta)$ foi a única selecionada em abril, porém não obteve menores valores dos critérios nos meses de fevereiro, março, outubro e novembro (Tabela 28).

Diante dos resultados dos critérios de informação, destaca-se então a distribuição $GaDTZ(k)$, pois em que 12 meses do ano tal distribuição obteve os maiores p valores em 8 meses, enquanto a $GeoTZ(q)$ atingiu os maiores p valores em maio, agosto e dezembro, já a $LiDTZ(\beta)$ apenas em abril.

Tabela 27 – Valores observados e esperados do tamanho das sequências de dias chuvosos da estação Porto São José do rio Paraná - 1985 a 2018.

Mês	Tamanho das sequências	Valor Observado	Valor Esperado LiDTZ	GeTZ	GaDTZ
Jan	1	89	87.41	89.15	88.90
	2	32	34.64	32.78	32.88
	3	13	12.53	12.05	12.14
	> 3	7	6.40	7.01	7.09
Fev	1	85	80.43	81.81	83.48
	2	25	29.81	28.27	27.24
	3	10	10.09	9.77	9.29
	> 3	5	4.67	5.16	4.99
Mar	1	76	72.58	73.02	71.89
	2	12	16.92	16.29	16.01
	3	3	3.59	3.63	4.28
	> 3	3	0.89	1.04	1.81
Abr	1	60	64.18	64.69	63.81
	2	24	17.29	16.58	16.26
	> 2	3	5.52	5.71	6.91
Mai	1	84	81.93	82.26	77.24
	2	14	15.75	15.25	17.19
	> 2	3	3.30	3.47	6.55
Jun	1	51	49.86	50.41	51.06
	2	14	15.32	14.61	13.79
	3	4	4.29	4.23	4.17
	> 3	2	1.53	1.73	1.98
Jul	1	56	54.18	54.57	55.42
	2	10	13.75	13.21	12.06
	3	5	3.18	3.19	3.19
	> 3	1	0.88	1.02	1.33
Ago	1	36	32.77	33.38	33.96
	2	7	12.61	11.95	11.65
	3	5	4.43	4.27	4.10
	> 3	4	2.17	2.39	2.28
Set	1	70	69.10	69.93	70.29
	2	20	22.06	21.03	20.20
	3	7	6.43	6.32	6.35
	> 3	3	2.41	2.72	3.14
Out	1	100	94.99	95.87	98.26
	2	21	26.86	25.70	23.35
	3	9	8.62	8.74	8.47
	> 3	1	0.51	0.67	0.90
Nov	1	92	88.31	89.13	90.91
	2	20	25.09	24.02	22.07
	3	6	6.51	6.47	6.24
	> 3	4	2.09	2.39	2.78
Dez	1	96	95.42	96.43	97.07
	2	28	28.85	27.55	26.22
	3	8	7.96	7.87	7.93
	> 3	3	2.77	3.15	3.77

Tabela 28 – Critérios de informação para as distribuições ajustadas as séries históricas mensais da estação Porto São José do rio Paraná - 1971 a 2018.

Mês	Critério de Informação	Distribuição		
		LiDTZ	GeTZ	GaDTZ
Jan	AIC	295.3801	295.3464	295.3445
	BIC	298.3289	298.2951	298.2932
Fev	AIC	249.7496	248.2556	247.8440
	BIC	252.5779	251.0839	250.6723
Mar	AIC	131.2315	130.4667	129.6851
	BIC	133.7748	133.0100	132.2284
Abr	AIC	134.1609	135.2088	137.1344
	BIC	136.6268	137.6748	139.6004
Mai	AIC	121.6083	120.9427	122.4341
	BIC	124.2234	123.5579	125.0492
Jun	AIC	122.6889	122.4303	122.3446
	BIC	124.9516	124.6930	124.6073
Jul	AIC	107.4173	107.1640	106.9363
	BIC	109.6940	109.4406	109.2130
Ago	AIC	108.1529	107.6683	107.5838
	BIC	110.1041	109.6195	109.5351
Set	AIC	176.8167	176.8763	177.1303
	BIC	179.4218	179.4815	179.7354
Out	AIC	211.4044	210.1471	208.7087
	BIC	214.2796	213.0223	211.5839
Nov	AIC	197.4566	196.6292	195.8277
	BIC	200.2607	199.4332	198.6317
Dez	AIC	228.1335	228.1459	228.5449
	BIC	231.0388	231.0512	231.4501

5.2.2 Estação Porto Rico

A análise exploratória mostrou que as séries mensais tiveram as menores amplitudes do tamanho das sequências dos dias chuvosos nos meses de abril e julho, com o maior máximo em fevereiro. Constatou-se que o mês mais chuvoso ocorreu em dezembro com 224 dias, seguido de janeiro com 206 dias de chuva. Notou-se também que o terceiro quartil com tamanho 2 ocorreu nos meses de janeiro, maio, setembro, outubro, dezembro, podendo constatar que 25% dos dias tiveram 2 dias de chuvas consecutivas devido a mediana ser 1 para todos os meses (Tabela 29).

Tabela 29 – Análise descritiva do tamanho das sequências de dias chuvosos das séries históricas mensais da estação do Porto Rico do Rio Paraná - 1971 a 2018.

Mês	n	Mo	Q_1	Md	$\overline{x}$	Q_3	Max	SD	CV (%)
Jan	316	1	1	1	1.53	2	7	0.99	64.91
Fev	269	1	1	1	1.40	1.25	8	0.89	64.11
Mar	198	1	1	1	1.28	1	4	0.56	44.20
Abr	164	1	1	1	1.21	1	3	0.47	39.44
Mai	199	1	1	1	1.36	2	4	0.69	50.88
Jun	157	1	1	1	1.26	1	4	0.63	50.46
Jul	115	1	1	1	1.16	1	3	0.39	34.14
Ago	111	1	1	1	1.32	1.25	4	0.64	48.66
Set	198	1	1	1	1.43	2	5	0.85	59.52
Out	260	1	1	1	1.38	2	5	0.76	55.05
Nov	255	1	1	1	1.34	1	5	0.69	51.55
Dez	310	1	1	1	1.38	2	5	0.73	53.22

Legenda: Número de dias chuvosos (n); moda (Mo); $1^{\underline{o}}$ quartil (Q_1); mediana (Md); média ($\overline{x}$); $3^{\underline{o}}$ quartil (Q_3); valor máximo (Max); desvio padrão (sd); coeficiente de variação (cv).

Destaca-se que a validade das suposições de que as séries mensais do tamanho das sequências de dias chuvosos da estação de Porto Rico mostraram-se independentes e identicamente distribuídas, segundo os testes Wald-Wolfowitz e Mann-Kendall, com nível de 1% de significância.

Assim, realizando o teste χ^2 de aderência para as distribuições $LiDTZ(\beta)$, $GeoTZ(q)$ e $GaDTZ(k)$, com nível de 1 % de significância, H_0 não foi rejeitada, logo evidenciando que as mesmas se adequaram ao tamanho das sequências em todas as séries mensais (Tabela 30).

Pode-se notar, na Figura 16, que em abril a distribuição $GaDTZ(k)$ não seguiu o perfil dos valores observados, isso pode ser constatado na Tabela 31, que o valor empírico no tamanho 1 foi de 110 e o valor esperado de 103.25. Fato semelhante, pode ser observado na Figura 17, uma vez que no mês de julho a mesma distribuição não acompanhou o valor observado, pois o valor empírico no tamanho 1 foi de 84 e o valor esperado de 76.86. Mas também, tal distribuição, desempenha uma boa alternativa para a predição dos períodos de chuva, veja o mês de maio em a distribuição $GaDTZ(k)$ apresentou o maior p valor, a

frequência absoluta do tamanho 1 foi 108 e o valor esperado mais próximo foi 108.18, no tamanho 2 o valor empírico foi 26 enquanto que o esperado mais próximo foi de 26.73.

Explorando o critério de informação pode-se notar que as distribuições truncadas avaliadas apresentaram valores de AIC e BIC muitos próximos, no entanto a distribuição $GaDTZ(k)$ apresentou os menores valores dos critérios avaliados nos meses de fevereiro e setembro, já as distribuições $LiDTZ(\beta)$ e $GeoTZ(q)$ tiveram os menores valores nos meses de março, abril e julho (Tabela 32).

Por efeito dos critérios de informação serem muitos próximos, ressalta que a distribuição $GaDTZ(k)$ apresentou os maiores p valores em 7 dos 12 meses anuais, com p valores aproximadamente iguais a distribuição $GeoTZ(q)$ em agosto e dezembro e $LiDTZ(\beta)$ com maiores valores em março, abril e julho.

Tabela 30 – Teste qui-quadrado de aderência para as distribuições ajustadas as séries históricas mensais da estação Porto Rico do rio Paraná - 1971 a 2018.

Mês	Estatística do teste	Distribuição		
		LiDTZ	GeTZ	GaDTZ
Jan	χ^2	4.2473	2.4932	2.1990
	p valor	0.2359	0.4765	0.5321
Fev	χ^2	4.0439	2.5804	1.0564
	p valor	0.1323	0.2752	0.5896
Mar	χ^2	0.2210	0.5226	1.7466
	p valor	0.8953	0.7700	0.4175
Abr	χ^2	0.7755	0.9299	3.9690
	p valor	0.6785	0.6281	0.1374
Mai	χ^2	0.9069	0.5111	0.3174
	p valor	0.6354	0.7744	0.8532
Jun	χ^2	6.7653	5.0643	3.0825
	p valor	0.0339	0.0794	0.2141
Jul	χ^2	0.6493	0.8429	5.0911
	p valor	0.4203	0.3585	0.0240
Ago	χ^2	1.7336	1.500	1.4717
	p valor	0.4202	0.4723	0.4791
Set	χ^2	4.7507	3.2115	2.0743
	p valor	0.0929	0.2007	0.3544
Out	χ^2	6.5555	5.1512	4.3023
	p valor	0.0874	0.1610	0.2306
Nov	χ^2	1.6972	1.2786	1.1162
	p valor	0.4280	0.5276	0.5722
Dez	χ^2	4.5324	3.8009	3.8001
	p valor	0.2094	0.2838	0.2839

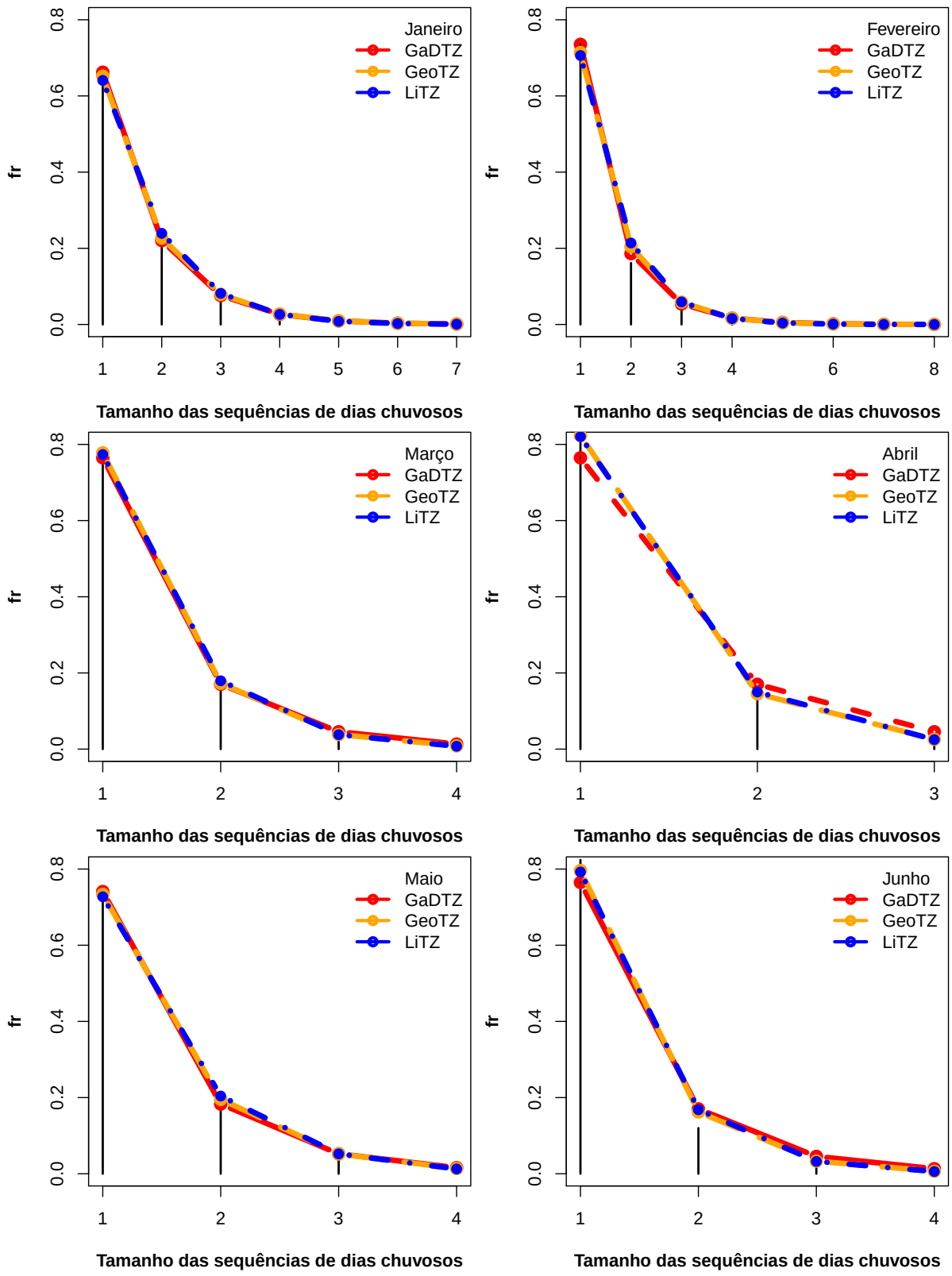

Figura 16 – Ajuste do tamanho das sequências de dias chuvosos das séries de janeiro a junho da estação Porto Rico do rio Paraná - 1971 a 2018.

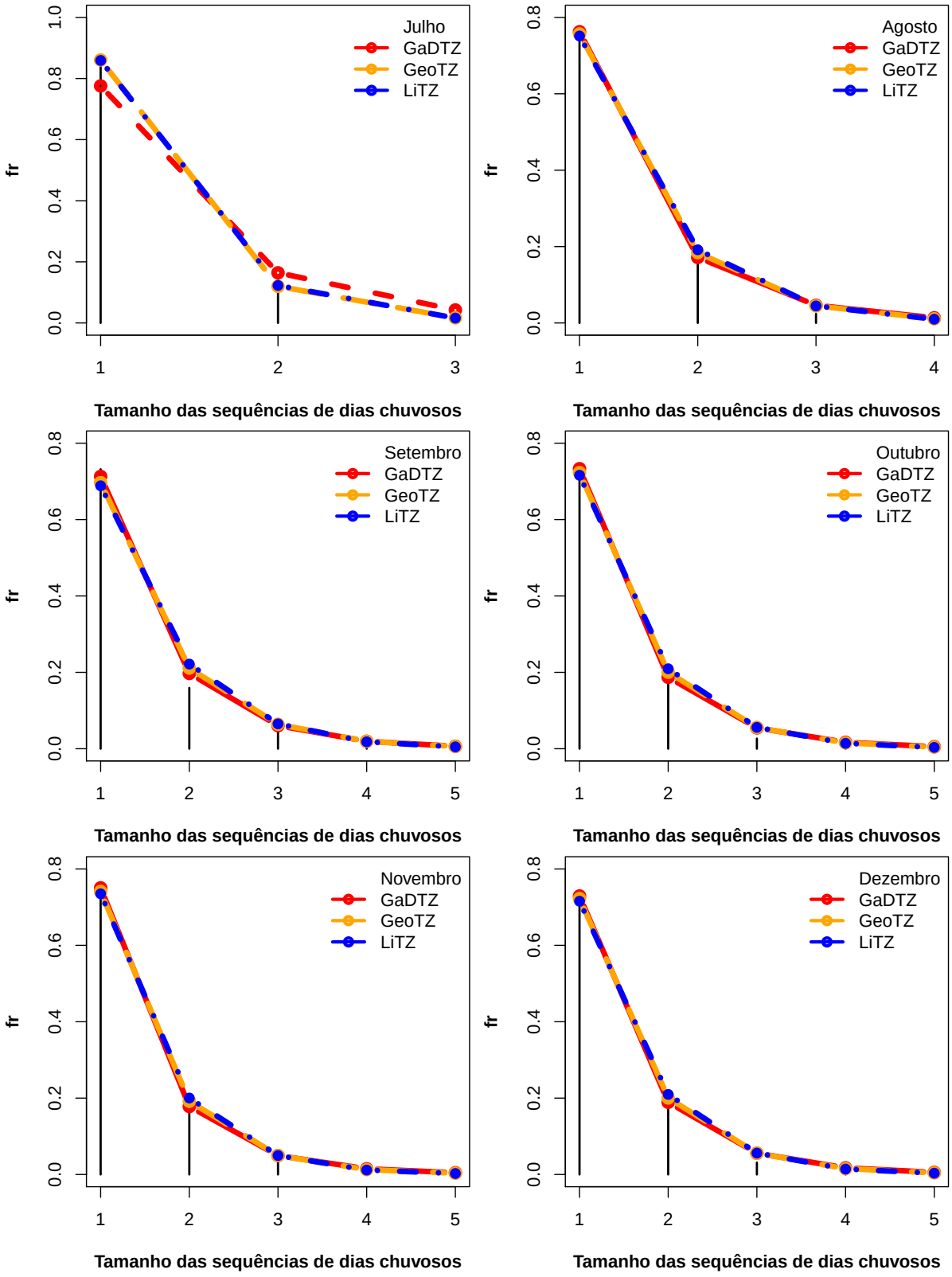

Figura 17 – Ajuste do tamanho das sequências de dias chuvosos das séries de julho a dezembro da estação Porto Rico do rio Paraná - 1971 a 2018.

Tabela 31 – Valores observados e esperados do tamanho das sequências de dias chuvosos da estação Porto Rico do rio Paraná - 1971 a 2018.

Mês	Tamanho das sequências	Valor Observado	Valor Esperado LiDTZ	GeTZ	GaDTZ
Jan	1	137	132.00	134.29	136.42
	2	47	49.31	46.74	45.39
	3	13	16.82	16.27	15.66
	4	4	5.44	5.66	5.50
	> 4	5	2.43	3.02	3.03
Fev	1	144	135.57	137.04	141.18
	2	31	41.08	39.22	35.71
	3	11	11.36	11.22	10.36
	> 3	6	3.97	4.50	4.73
Mar	1	118	119.08	119.78	117.79
	2	29	27.62	26.62	26.22
	3	6	5.84	5.91	7.03
	> 3	1	1.45	1.69	2.97
Abr	1	110	110.74	111.12	103.25
	2	21	20.25	19.65	22.98
	> 2	4	4.00	4.22	8.75
Mai	1	108	106.17	107.11	108.18
	2	26	29.80	28.53	26.73
	3	9	7.63	7.59	7.64
	> 3	3	2.41	2.76	3.44
Jun	1	103	99.03	99.52	95.60
	2	15	21.01	20.28	21.28
	> 2	4	4.00	4.22	8.75
Jul	1	84	85.05	85.22	76.81
	2	14	12.14	11.85	16.24
	> 2	1	1.79	1.91	5.93
Ago	1	63	63.11	63.57	64.05
	2	17	16.09	15.46	14.41
	3	2	3.74	3.76	3.89
	> 3	2	1.05	1.21	1.65
Set	1	101	95.01	96.18	98.39
	2	22	30.59	29.15	27.21
	3	9	8.99	8.83	8.35
	> 3	6	3.41	3.84	4.04
Out	1	137	134.61	135.94	137.86
	2	39	39.38	37.64	35.17
	3	5	10.51	10.42	10.26
	4	5	2.66	2.89	3.18
	> 4	2	0.84	1.11	1.53
Nov	1	142	138.92	140.08	141.86
	2	32	37.83	36.26	33.65
	3	12	9.40	9.38	9.38
	> 3	3	2.86	3.28	4.12
Dez	1	161	160.28	161.86	163.36
	2	49	46.97	44.90	42.36
	3	7	12.55	12.46	12.49
	4	5	3.18	3.45	3.91
	> 4	2	1.02	1.33	1.89

Tabela 32 – Critérios de informação para as distribuições ajustadas as séries históricas mensais da estação Porto Rico do rio Paraná - 1971 a 2018.

Mês	Critério de Informação	Distribuição LiDTZ	GeTZ	GaDTZ
Jan	AIC	412.4708	410.4384	410.0047
	BIC	415.7986	413.7663	413.3326
Fev	AIC	326.7997	324.1305	321.5391
	BIC	330.0572	327.3880	324.7965
Mar	AIC	211.3594	211.7637	213.6359
	BIC	214.3964	214.8006	216.6729
Abr	AIC	154.7537	155.0288	160.0310
	BIC	157.6590	157.9341	162.9363
Mai	AIC	232.8027	232.6712	232.9402
	BIC	235.7863	235.6549	235.9238
Jun	AIC	161.6460	160.7757	160.9585
	BIC	164.4743	163.6040	163.7868
Jul	AIC	94.5286	94.7778	101.5435
	BIC	97.1237	97.3729	104.1386
Ago	AIC	125.1319	125.1633	125.5984
	BIC	127.5627	127.5941	128.0293
Set	AIC	246.2293	244.9104	243.9221
	BIC	249.1565	247.8376	246.8493
Out	AIC	309.5826	308.8123	308.4732
	BIC	312.8190	312.0488	311.7097
Nov	AIC	294.0755	293.6296	293.5857
	BIC	297.3173	296.8714	296.8275
Dez	AIC	368.4202	368.1097	368.4999
	BIC	371.8318	371.5213	371.9115

Capítulo 6

Considerações Finais

Diante dos resultados obtidos pode-se declarar que flexibilidade da função massa de probabilidade da distribuição $GaGDTZ\ (k, \theta, c)$ foi eficiente para ajustar todas as séries históricas anuais do tamanho das sequências de dias chuvosos das estações de Porto São José, Porto Rico do rio Paraná e capitais do Brasil, uma vez que a distribuição mostrou um bom desempenho nos ajustes dos dados, sendo que o Brasil apresenta uma grande variabilidade climática e sete dos nove tipos de clima, classificados por Dubreuil et al. (2018), que existem no Brasil estão presentes nas capitais.

Ressalta-se que as estimativas intervalares dos parâmetros foram essenciais para verificar se algumas das distribuições dos casos especiais fornecem ajustes adequados para o tamanho das sequências de dias chuvosos, o qual foi possível sinalizar que as distribuições truncadas em zero como Rayleigh discreta, Rayleigh generalizada discreta, hydrograph discreta, Maxwell-Boltzmann discreta e half-normal discreta não foram adequadas nos ajustes dos dados desse estudo.

Por sua vez, as distribuições $GaDTZ(k, \theta)$ e $WeDTZ(c, q)$ mostraram ter os melhores comportamentos segundo os critérios de informação para maioria das capitais na Região Norte, Nordeste, Centro-Oeste e Sudeste. Enquanto que, a distribuição $LNDTZ(\mu, \sigma)$ foi mais adequada nos ajustes do tamanho das sequências de dias chuvosos das estações de Manaus, Fortaleza, São Paulo e Porto Alegre. Acrescenta-se também que a estação de Campo Grande apresentou ajuste mais adequado com a distribuição $GeoTZ(q)$, assim como para as estações de Curitiba e Florianópolis juntamente com a distribuição $GaDTZ(k)$, uma vez que essas estações apresentaram amplitude do tamanho das sequências de dias chuvosos inferior a 13 dias.

Destaca-se, também, que foi importante avaliar o desempenho das distribuições $GaDTZ(k)$, $GeoTZ(q)$ e $LiDTZ(\beta)$ para ajustar o tamanho das sequências dos dias chuvosos mensais das estações de Porto São José e Porto Rico, pois todas essas distribuições apresentaram uma boa adequabilidade no ajuste e valores dos critérios de informação muito próximos para a maioria dos meses, porém constatou-se que a distribuição $GaDTZ(k)$ apresentou os maiores p valores para a maioria das 24 séries avaliadas, ou seja, 8 meses para estação de Porto São José e 7 meses para estação de Porto Rico. Inclusive foi possível constatar que as distribuições $GaDTZ(k)$ e $GeoTZ(q)$ dependem da amplitude curta do tamanho das sequências dias chuvosos e dos valores esperados próximos dos empíricos, como foram observados nas estações de Porto São José e Porto Rico.

Logo, os resultados mostraram que utilizar distribuição $GaGDTZ(k, \theta, c)$ é vantajosa, visto que ela ainda pode indicar se algumas das 10 distribuições dos casos particulares fornecem ajustes adequados.

Capítulo 7

Referências Bibliográficas

Abouammoh, A.M. ; Alhazzani, N.S. On Discrete Gamma Distribution. Communications in Statistics - Theory and Methods, v.44, n.14, p.3087-3098, 2015.

Achcar, J. A.; Ramos, P. L.; Martinez, E. Z. Some computational aspects to find accurate estimates for the parameters of the generalized gamma distribution. *Pesquisa Operacional*, v.37, n.2, p.365-385, 2017.

Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, v.19, n.6, p.716-723, 1974.

Alamatsaz, M.H. ; Dey, S.; Dey, T.; Shams Harandi, S. Discrete generalized Rayleigh distribution. Pak. J. Statist., v.32, n.1, p.1-20, 2016.

ANA - Agência Nacional de Águas (Brasil). Medição de descarga líquida em grandes rios: manual técnico. 2º ed., Brasília, 94p., 2014.

Amin, M.T.; Rizwan, M.; Alazba, A.A. A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan. *Open Life Sci*, n.11, v.1, p.432-440, 2016.

Bakouch, H. S.; Jazi, M. A.; Nadarajah, S. A new discrete distribution. *Statistics*, v.48, n.1, p.200-240, 2012.

Bi, Z.; Faloutsos, C.; Korn, F. The "DGX" distribution for mining massive skewed data. 7º Conf. Knowledge discovery and data Mining, San Francisco, 2001.

Boswell, M.T.; Ord, J.K; Patil, G.P. Chance mechanisms underlying univariate distributions. *Statistical distributions in ecological work*, v.3, p. 1-156, 1979.

Caloiero, T.; Coscarelli, R.; Ferrari, E. e Sirangelo, B. Analysis of dry spells in Southern Italy (Calabria). *Water*, v.7, n.12, p.3009-3023, 2015.

Chakraborty, S. A new discrete distribution related to generalized gamma distribution and its properties. *Communications in Statistics - Theory and Methods*, v.44, n.8, p.1691-1705, 2015.

Chakraborty, S; Chakravarty, D. A discrete Gumbel distribution. *arXiv:1410.7568 [math.ST]*, Outubro, 2014.

Chakraborty, S.; Chakravarty, D. Discrete gamma distributions: properties and parameter estimation. *Commun. Stat. Theory. Methods.*, v.41, n.18, p.3301-3324, 2012.

Chowdhury, R.K. e Beecham, S. Characterization of rainfall spells for urban water management. *International Journal of Climatology*, v. 33, n.4, p. 959-967, 2013.

Cremoneze, I.Z.; Peralta, D.; Mazucheli, J.; Emanuelli, I.P. Análise de frequências das precipitações máximas mensais observadas nos estados do Paraná e Rio Grande Do Sul. *Enciclopédia Biosfera*, v.14 n.25; p. 48-57, 2017.

Davison, A. C. **Statistical Models**. Cambridge: Cambridge University Press, New York. 2011.

Delignette-Muller, M.L.; Dutang, C. fitdistrplus: An R package for fitting distributions. Journal of Statistical Software, v.64, n.4, p.1-34, 2015.

Deni, S.M.; Jemain, A.A.; Ibrahim, K. The spatial distribution of wet and dry spells over Peninsular Malaysia. *Theoretical and Applied Climatology*, v.94, p.163-173, 2008.

Deni, S.M.; Jemain, A.A. Fitting the distribution of dry and wet spells with alternative probability models. *Meteorology and Atmospheric Physics*, v.104, p.13-27, 2009.

Deni, S.M.; Jemain, A.A.; Ibrahim, K. The best probability models for dry and wet spells in Peninsular Malaysia during monsoon seasons. *International Journal of Climatology*, v.30, p.1194-1205, 2010.

Devi, B.; Gupta, R. e Kumar, P. Sequential analysis of zero truncated geometric distribution. *J. Stat. Appl. Pro.*, v.6, n.2, p.385-389, 2017.

Dubreuil, V.; Fante, K.P.; Planchon, O.; Sant'anna Neto, J.L. Os tipos de climas anuais no Brasil : uma aplicação da classificação de Köppen de 1961 a 2015. *Revista franco-brasilera de geografia*, v.37, n.37, 2018.

Gómez-Déniz, E.; Calderin-Ojeda, E. The discrete Lindley distribution: properties and applications. *J. Statist. Computat. Simul.*, v.81, n.11, p.1405-1416, 2011.

Gouvea, R. L.; Campos, C.C.; Menezes, J.T.; Moreira, G.F. Análise de frequência de precipitação e caracterização de anos secos e chuvosos para a bacia do rio Itajaí. Rev. Brasileira de Climatologia, v. 22, 2018.

Greene, W.H. **Econometric Analysis**. Prentice Hall, New Jersey, 7º Edição. 2012.

Hasanshahi, Z; Azmi, P.; Gholizadeh, M.H.; Khajezadeh, M. The flexibility of the generalized gamma distribution in modeling the fading based on Kullback-Leibler and Kolmogorov-Smirnov criteria. *IEEE Access*, v.8, p.8393-8404, 2020.

Husak, G. J.; Michaelsen, J.; Funk, C. Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. *International Journal of Climatology*, v.27, 2007.

J, K.; Ngesa, O.; Orwa, G. On Generalized Gamma Distribution and Its Application to Survival Data. *International Journal of Statistics and Probability*, v. 8, n. 5, p. 85-102, 2019.

Johnson, N.L.; Kemp, A.W.; Kotz, S. **Univariate Discrete Distributions**. 3º Edição. John Wiley & Sons. 653 p, 2005.

Kemp, A.W. The discrete half-normal distribution. *Advances in Mathematical and Statistical Modeling Statistics for Industry and Technology*, p.353-360, 2008.

Klugman, S.A.; Panjer, H.H.; Willmot, G.E. **Loss Models: From Data to Decisions**. 4º Edição. Wiley. 536p., 2012.

Krishna, H., Pundir, P. S. Discrete Burr and discrete Pareto distributions. *Statist. Methodol.*, v.6, p.177-188, 2009.

Krishna, H.; Pundir, P. S. Discrete Maxwell distribution. *InterStat*, v.3, n.11, 2007.

Kumar, V.; Shanu; Jahangeer. Statistical distribution of rainfall in Uttarakhand, India. *Applied Water Science*, n.7, v.8, p.4765-4776, 2017.

Li, Z.; Li, Y.; Shi, X.; Li, J. The characteristics of wet and dry spells for the diverse climate in China. Global and Planetary Change, v.149, p.14-19, 2017.

Mandapaka, P.V.; Qin, X.; Yat-Man Lo, E. Seasonal and interannual variability of wet and dry spells over two urban regions in the western maritime continent. Journal of Hydrometeorology, v.17, n.5, p.1579-1600, 2016.

Matheson, M.; Muñoz, A.; Cox, C. Describing the flexibility of the generalized gamma and related distributions. *Journal of Statistical Distributions and Applications*, v.15, n.4, 2017.

Mathugama, S.C.; Peiris, T.S.G. Critical evaluation of dry spell research. *International Journal of Basic & Applied Sciences*, v.11, n.6, p. 153-160, 2011.

Mazucheli, J.; Emanuelli, I. P. Aplicação da distribuição Nakagami na análise de dados de precipitação. *Revista Brasileira de Meteorologia*, v.34, n.1, 1-7, 2019.

Moreira, P.S.P.; Dallacort, R.; Magalhães, R.A.; Inoue, M.H.; Stieler, M.C.; Silva, D.J.; Martins, J.A. Distribuição e probabilidade de ocorrência de chuvas no município de Nova Maringá-MT. Revista de Ciências Agro-Ambientais, v.8, n.1, p.9-20, 2010.

Mzezewa, J.; Misi, T.; Van Rensburg, L. Characterisation of rainfall at a semi-arid ecotope in the Limpopo Province (South Africa) and its implications for sustainable crop production. *Water SA* , v.36, n.1, 2010.

Nadarajah, H. S. e Gupta, A. K. A generalized gamma distribuition with application to drought data. *Mathematics and Computers in Simulation*, v.74, n.1, p.1-7, 2007.

Nakagawa T.; Osaki S. The Discrete Weibull Distribution. *IEEE. Transactions on Reliability*, v.R-24, n.5, p.300-301. 1975.

Oliveira, R.P. Aplicações de alguns métodos de discretização nas distribuições Lindley contínuas. Dissertação (Mestrado) - Programa de Pós-Graduação em Bioestatística, Universidade Estadual de Maringá, Maringá - PR, 2016.

Pal, S.; Yu, H.; Loucks, Z. D.; Harris, I. M. Illustration of the flexibility of generalized gamma distribution in modeling right censored survival data: Analysis of two cancer datasets. Annals of Data Science. 2019.

Papalexiou, S.M. e Koutsoyiannis, D. A global survey on the seasonal variation of the marginal distribution of daily precipitation. Advances in Water Resources, n. 94, p. 131-145, 2016.

Para, B.A. e Jan, T.R. Discrete version of log-logistic distribution and its applications in genetics. *Int. J. Modern Math. Sci.*, v.14, n.4, p.407-422, 2016.

Pohlert T. trend: Non-parametric trend tests and change-point detection. R package version 1.1.1. 2018.

Ramos, P.L.; Moala, F.A. A aplicação da distribuição exponencial geométrica estendida para modelagem de dados pluviométricos. *Rev. bras. meteorol.*, v. 29, n. 4, p. 613-620, 2014.

Rayner, D.; Achberger, C.; Chen, D. A multi-state weather generator for daily precipitation for the Torne River basin, northern Sweden/western Finland. Advances in Climate Change Research, v.7, n.1-2, p.70–81, 2016.

Reis, T.C.S. Extensões da distribuição Weibull aplicadas na análise de séries climatológicas. Dissertação (Mestrado) - Programa de Pós-Graduação em Matemática Aplicada e Computacional, Faculdade de Ciências e Tecnologia de Presidente Prudente, Presidente Prudente-SP, 2017.

Roy, D. Discrete Rayleigh distribution. *IEEE Trans. Reliab.* v.53, n.2, p.255-260, 2004.

Santos, J. ; Ferreira, V. A variabilidade pluviométrica na Mesorregião do Triângulo Mineiro/Alto Paranaíba-MG. *GeoTextos*, v. 12, n. 1, p. 233-265, 2016.

Schwarz, G. Estimating the dimensional of a model. *Annals of Statistics*, v.6, n.2, p.461-464, 1978.

Şen, Z. **Applied Drought Modeling, Prediction, and Mitigation**. Elsevier, Amsterdam, 484p., 2015.

Silva, V.P.R.; Pereira, E.R.R.; Azevedo, P.V.; Sousa, F.A.S.; Sousa, I.F. Análise da pluviometria e dias chuvosos na região Nordeste do Brasil. Revista Brasileira de Engenharia Agrícola e Ambiental, v.15, n.2, p.131-138, 2011.

Shanker, R.; Shukla, K. K. Zero-truncated Poisson-Garima distribution and its applications. *Biostat Biometrics Open Acc J*, v.3, n.1, p.1-6, 2017.

Stacy, E. W. A generalization of the gamma distribution. *Ann. Math. Statist.*, v.33, p.1187-1192, 1962.

Sukla, M.K.; Mangaraj, A.K.; Sahoo, L.N.; Sethy, K.M. A comparative study of three models for the distribution of wet and dry spells in the Mahanadi delta. *New York Science Journal*, v.5, n. 11, p.54-61, 2012.

Tian, G.L.; Ding, X.; Liu, Y.; Tang, M.-L. Some new statistical methods for a class of zero-truncated discrete distributions with applications. *Computational Statistics*, 2019.

Thomasson, R.L.; Kapadia, C.H. On estimating the parameter of a truncated geometric distribution. *Annals of the Institute of Statistical Mathematics*, v.20, n.1, p.519-523, 1968.

Thurai, M., e Bringi, V. N. Application of the generalized gamma model to represent the full rain drop size distribution spectra. *J. Appl. Meteoro. Climatol.*, v.57, n.5, p.1197-1210, 2018.

Vela, R. H. N.; Dallacort, R.; Nied, A. H. Distribuição decendial, mensal e totais de precipitação na região de Tangará da Serra - MT. Anais Congresso Brasileiro de Engenharia Agrícola, SBEA, Bonito, p.1-4, 2007.

Wald, A.; Wolfowitz, J. An exact test for randomness in the non-parametric case based on serial correlation. *Annals of Mathematical Statistics*, v.14, p.378-388, 1943.

WMO. World Meteorological Organization. **Guide to Hydrological Practices**. Management of Water Resources and Application of Hidrological Practices. Volume II, 300p., 2009.

Ye, L.; Hanson, L.S.; Ding, P.; Wang, D.; Vogel, R.M. The probability distribution of daily precipitation at the point and catchment scales in the United States. *Hydrol. Earth Syst. Sci. Discuss.*, v.22, n.12, p.6519-6531, 2018.

Yee, T. W. **Vector generalized linear and additive models**. Springer-Verlag, New York. 589p., 2015.

Zografos, K.; Balakrishnan, N. On families of beta and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, v.6, n.4, p. 344-362, 2009.

Zolina, O.; Simmer, C.; Belyaev, K.; Gulev, S. K.; Koltermann, P. Changes in the duration of European wet and dry spells during the last 60 years. *Journal of Climate*, v.26, n.6, p. 2022-2047, 2013.

Zuur, A.F.; Ieno, E.N.; Walker, N.J.; Saveliev, A.A.; Smith, G.M. Zero-truncated and zero-inflated models for count data. In: Mixed effects models and extensions in ecology with R. Statistics for Biology and Health. Springer, New York, p.261-293, 2009.