

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM SANEAMENTO,**  
**MEIO AMBIENTE E RECURSOS HÍDRICOS**

**GERAÇÃO ESTOCÁSTICA DE PRECIPITAÇÃO**  
**SUBDIÁRIA E SIMULAÇÃO DE EVENTOS**  
**EXTREMOS**

**Milena Guerra de Aguiar**

**Belo Horizonte**

**2019**

**GERAÇÃO ESTOCÁSTICA DE PRECIPITAÇÃO  
SUBDIÁRIA E SIMULAÇÃO DE EVENTOS  
EXTREMOS**

**Milena Guerra de Aguilar**

**Milena Guerra de Aguiar**

# **GERAÇÃO ESTOCÁSTICA DE PRECIPITAÇÃO SUBDIÁRIA E SIMULAÇÃO DE EVENTOS EXTREMOS**

Dissertação apresentada ao Programa de Pós-graduação em Saneamento, Meio Ambiente e Recursos Hídricos da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Saneamento, Meio Ambiente e Recursos Hídricos.

Área de concentração: Recursos Hídricos

Linha de pesquisa: Modelagem de Processos Hidrológicos

Orientador: Veber Afonso Figueiredo Costa

Belo Horizonte

Escola de Engenharia da UFMG

2019

A283g

Aguilar, Milena Guerra de.

Geração estocástica de precipitação subdiária e simulação de eventos extremos [recurso eletrônico] / Milena Guerra de Aguilar.- 2019.  
xiii, 121 f., enc.: il.

Orientador: Veber Afonso Figueiredo Costa.

Mestrado (dissertação) - Universidade Federal de Minas Gerais,  
Escola de Engenharia.

Anexos: f. 104-121.

Bibliografia: f. 97-103.

1. Engenharia sanitária - Teses. 2. Recursos hídricos -  
Desenvolvimento - Teses. 3. Precipitação (Meteorologia) - Teses.  
I. Costa, Veber Afonso Figueiredo. II. Universidade Federal de Minas  
Gerais, Escola de Engenharia. III. Título.

CDU: 628(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS

Escola de Engenharia

*Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos*

*Avenida Antônio Carlos, 6627 - 4º andar - 31270-901 - Belo Horizonte – BRASIL*

*Telefax: 55 (31) 3409-1882 - posgrad@desa.ufmg.br*

*http://www.smarh.eng.ufmg.br*

## FOLHA DE APROVAÇÃO

Geração Estocástica de Precipitação Subdiária e Simulação de Eventos Extremos

**MILENA GUERRA DE AGUILAR**

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

*M. Afonso Figueiredo Costa*

Prof. VEBER AFONSO FIGUEIREDO COSTA

*Francisco Eustáquio Oliveira e Silva*

Prof. FRANCISCO EUSTÁQUIO OLIVEIRA E SILVA

*Luz Rafael Palmier*

Prof. LUZ RAFAEL PALMIER

*Carlos Rogério de Mello*

Prof. CARLOS ROGÉRIO DE MELLO

Aprovada pelo Colegiado do PG SMARH

*Antonio Teixeira de Matos*

Prof. Antonio Teixeira de Matos  
Coordenador

Versão Final aprovada por

*M. Afonso Figueiredo Costa*

Prof. Veber Afonso Figueiredo Costa  
Orientador

Belo Horizonte, 18 de junho de 2019.

## AGRADECIMENTOS

Agradeço aos meus pais e meus irmãos, pelo incentivo e apoio durante a elaboração desta dissertação.

Ao meu orientador, professor Veber Costa, pela constante dedicação, colaboração e suporte no decorrer deste trabalho.

Ao professor Wilson Fernandes, pela ajuda no algoritmo diário e contribuições ao longo desta pesquisa.

Ao professor Éber Andrade, pela disponibilização dos dados para o desenvolvimento deste trabalho.

Aos amigos da Eng. Civil, em especial Paulo Alfenas e Roberto Rangel, pelo imenso apoio e incentivo durante esta trajetória.

Aos amigos da Galerinha, pelo alívio emocional, tantos anos de amizade que já os considero como família.

Obrigada a todos, sem vocês essa jornada também não seria possível.

*“A winner is a dreamer who never gives up”*

*Nelson Mandela*

## RESUMO

Séries de precipitação de alta resolução temporal são necessárias para a modelagem de eventos regulares e extremos em projetos de sistemas pluviais urbanos, simulação de vazões em pequenas bacias, estudos de balanços hídricos e modelagem de inundações de curta duração, por exemplo. Não obstante, séries longas e contínuas de precipitação subdiária são, via de regra, difíceis de ser encontradas. De modo a obter séries compatíveis com as premissas mencionadas, são amplamente utilizados os geradores estocásticos de precipitação, que buscam reproduzir as propriedades estatísticas das séries de precipitação observadas, além de quantificar as incertezas e avaliar o risco envolvido nas estimativas dos quantis raros e extremos de chuva, para um conjunto adequado de durações.

Nesta pesquisa foi desenvolvido um gerador estocástico de precipitação para a escala subdiária, apto a simular de maneira apropriada tanto os eventos regulares quanto aqueles mais extremos. Para isso, utilizou-se de um gerador estocástico de precipitação diária acoplado a um gerador de precipitação subdiária. O emprego do modelo diário buscou introduzir variabilidade às séries a serem desagregadas, obtendo estimativas mais confiáveis daqueles eventos de precipitação com reduzida probabilidade de superação. Após a geração das séries diárias, foi empregado um método não paramétrico de reamostragem juntamente com uma abordagem por similaridade regional, onde os “fragmentos” de precipitação subdiária são aleatoriamente amostrados de pluviógrafos nas proximidades, condicionados à altura de chuva diária no local de interesse.

O método proposto foi aplicado para um grupo de 40 postos pluviográficos, realizando a desagregação da precipitação diária para as durações de 60, 180, 360 e 720 minutos. Os resultados indicaram um desempenho apropriado tanto para a geração da precipitação diária quanto para a sua desagregação, reproduzindo as estatísticas mensais, diárias e subdiárias, para as durações analisadas. Adicionalmente, o comportamento dos máximos anuais, mesmo para baixas probabilidades de excedência, foi relativamente bem descrito, abrangendo a variabilidade esperada dos quantis. De forma geral, a abordagem proposta mostrou-se uma alternativa coerente para simular séries contínuas de precipitação subdiária a partir de registros de menor resolução temporal.

## ABSTRACT

Rainfall time series with high temporal resolution are often required for estimating storm events for the design of urban drainage systems, for performing rainfall-runoff simulation in small catchments and for modeling flash-floods. Nonetheless, large and continuous sub-daily rainfall samples are, most often than not, unavailable. In order to obtain time series with the mentioned premises, the stochastic rainfall generators are widely used. Those generators aim to reproduce the statistical proprieties of the observed rainfall series, besides quantify the uncertainties and assess the risks involved in the extreme rainfall quantiles estimative, for an appropriate group of durations.

In this research, a sub-daily stochastic rainfall generator was developed. This generator is able to appropriate simulate both extreme and regular events. To achieve this, a daily and sub-daily stochastic rainfall generators were coupled. The use of a daily model allowed to introduce variability to the series to be disaggregated, obtaining more reliable estimates for the low exceedance probability rainfall events. After generating the daily series, a non-parametric approach of resampling was used herewith a regionalised similarity approach, where the “fragments” of subdaily rainfall are randomly sampled of sub-daily record gauges at nearby stations, conditioned on the daily precipitation amount at the location of interest.

The proposed disaggregation method was applied to a set of 40 rainfall gauging stations. Disaggregation of daily rainfall was performed for the durations of 60, 180, 360 and 720 minutes. Results indicated an appropriate performance for the daily rainfall generation as well as for its disaggregation, reasonably reproducing monthly, daily and sub-daily summary statistics, for the evaluated durations. In addition, the annual block-maxima behavior, even for low exceedance probabilities, was relatively well described, properly summarizing the expected variability in the quantiles. Overall, the proposed approach proved a sound alternative for simulating continuous sub-daily rainfall amounts from coarse-resolution records.



# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
<b>2</b>	<b>OBJETIVOS</b>	<b>4</b>
2.1	OBJETIVO GERAL	4
2.2	OBJETIVOS ESPECÍFICOS	4
<b>3</b>	<b>REVISÃO DA LITERATURA</b>	<b>5</b>
3.1	GERAÇÃO ESTOCÁSTICA DE PRECIPITAÇÃO DIÁRIA	5
3.1.1	<i>Introdução</i>	5
3.1.2	<i>Modelos bipartidos</i>	6
3.1.3	<i>Modelos não paramétricos e semi-paramétricos</i>	7
3.1.4	<i>Modelos de matriz de probabilidade de transição - MPT</i>	8
3.1.5	<i>Modelos híbridos</i>	9
3.2	LIMITE SUPERIOR DE PRECIPITAÇÕES	13
3.3	TEORIA BAYESIANA PARA ESTIMAÇÃO DE PARÂMETROS	16
3.4	GERAÇÃO ESTOCÁSTICA DE PRECIPITAÇÃO SUBDIÁRIA	20
3.4.1	<i>Introdução</i>	20
3.4.2	<i>Modelos baseados em Processos de Poisson</i>	22
3.4.3	<i>Modelos baseados em invariância de escala</i>	24
3.4.4	<i>Modelos paramétricos</i>	26
3.4.5	<i>Modelos não-paramétricos</i>	27
<b>4</b>	<b>METODOLOGIA</b>	<b>32</b>
4.1	ETAPA 1 – DADOS HIDROLÓGICOS	32
4.2	ETAPA 2 – GERADOR ESTOCÁSTICO DE PRECIPITAÇÃO DIÁRIA	35
4.2.1	<i>Construção das distribuições a priori dos parâmetros do modelo LN4</i>	39
4.2.2	<i>Construção das distribuições a posteriori dos parâmetros do modelo LN4</i>	42
4.2.3	<i>Avaliação da eficiência do modelo diário</i>	43
4.3	ETAPA 3 – GERAÇÃO DOS DADOS EM ESCALA SUBDIÁRIA	44
4.3.1	<i>Definição da similaridade hidrológica</i>	45
4.3.2	<i>Definição da similaridade fisiográfica</i>	46
4.3.3	<i>Algoritmo de desagregação da chuva diária em subdiária</i>	48
4.4	ETAPA 4 – VALIDAÇÃO E AVALIAÇÃO DA EFICIÊNCIA DO MODELO	51
<b>5</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>53</b>
5.1	AVALIAÇÃO DOS DADOS HIDROLÓGICOS	53
5.2	GERADOR ESTOCÁSTICO DIÁRIO	56
5.2.1	<i>Distribuição a priori para o limite superior</i>	56
5.2.2	<i>Distribuição a posteriori para os parâmetros da distribuição LN4</i>	56
5.2.3	<i>Definição do limiar entre os eventos regulares e extremos</i>	59
5.2.4	<i>Avaliação do desempenho do gerador estocástico diário</i>	63
5.3	GERADOR ESTOCÁSTICO SUBDIÁRIO	65
5.3.1	<i>Definição da similaridade hidrológica e fisiográfica</i>	65
5.3.2	<i>Calibração do modelo de desagregação</i>	71
5.4	VALIDAÇÃO E AVALIAÇÃO DA EFICIÊNCIA DO MODELO	76
<b>6</b>	<b>CONCLUSÕES</b>	<b>92</b>
<b>7</b>	<b>RECOMENDAÇÕES</b>	<b>96</b>
	<b>REFERÊNCIAS</b>	<b>97</b>
	<b>APÊNDICE I</b>	<b>104</b>

## LISTA DE FIGURAS

Figura 3.1 – Resumo dos principais modelos de geração de chuva diária .....	5
Figura 3.2 – Efeito de cada parâmetro na distribuição LN4.....	13
Figura 3.3 – Estimativa do fator $k_m$ para o método estatístico de estimação da PMP .....	15
Figura 3.4 – Resumo da teoria Bayesiana .....	19
Figura 3.5 – Resumo dos principais modelos de geração de chuva subdiária.....	22
Figura 4.1 – Localização das estações pluviográficas utilizadas neste trabalho .....	34
Figura 4.2 – Procedimento para definição do estado do primeiro dia de simulação.....	36
Figura 4.3 – Procedimento para definição do estado a partir do segundo dia de simulação ....	37
Figura 4.4 – Histograma de frequências das estimativas de PMP de 1 dia em Minas Gerais..	40
Figura 4.5 – Resumo da metodologia do gerador diário .....	43
Figura 4.6 – Resumo da metodologia do gerador subdiário.....	51
Figura 5.1 – Comparação da série temporal de precipitação diária do SNIRH e a série subdiária agregada do pluviógrafo para a estação de Caeté.....	55
Figura 5.2 – Variação dos valores dos parâmetros do modelo LN4 ao longo da simulação....	57
Figura 5.3 – Densidade e histograma dos valores dos parâmetros do modelo LN4 ao longo da simulação .....	58
Figura 5.4 – Curvas de quantis de precipitações diárias máximas anuais para diferentes limiares entre eventos regulares e extremos .....	60
Figura 5.5 – Precipitações médias mensais para diferentes limiares entre eventos regulares e extremos (continua).....	61
Figura 5.6 – Comparação das estatísticas entre as séries simuladas (preto) e as séries observadas (azul).....	63
Figura 5.7 – Comparação das médias e coeficiente de variação mensal e anual entre as séries simuladas (cinza) e as séries observadas (preto) .....	65
Figura 5.8 – Resultados das métricas obtidas para a duração de 60 min .....	76
Figura 5.9 – Resultados das métricas obtidas para a duração de 180 min .....	77
Figura 5.10 – Resultados das métricas obtidas para a duração de 360 min .....	77
Figura 5.11 – Resultados das métricas obtidas para a duração de 720 min .....	78
Figura 5.12 – Resultados do desvio das métricas obtidas para a duração de 60 min.....	80

Figura 5.13 – Resultados do desvio das métricas obtidas para a duração de 180 min .....	80
Figura 5.14 – Resultados do desvio das métricas obtidas para a duração de 360 min .....	81
Figura 5.15 – Resultados do desvio das métricas obtidas para a duração de 720 min .....	81
Figura 5.16 – Curvas de quantis das precipitações máximas anuais para a estação de Caeté para todas as durações .....	82
Figura 5.17 – Resultados do desvio das métricas obtidas para a duração de 60 min do modelo unicamente subdiário .....	85
Figura 5.18 – Curvas de quantis das precipitações máximas para a duração de 60 minutos ...	85
Figura 5.19 – Probabilidade de similaridade em função da latitude para cada estação do ano	86
Figura 5.20 – Probabilidade de similaridade em função da longitude para cada estação do ano .....	87
Figura 5.21 – Probabilidade de similaridade em função da multiplicação da diferença de longitude e latitude para cada estação do ano.....	87
Figura 5.22 – Probabilidade de similaridade em função da elevação para cada estação do ano .....	88
Figura 5.23 – Estações utilizadas para desagregação – Duração 60 minutos – Estações do ano Verão e Outono.....	90
Figura 5.24 – Estações utilizadas para desagregação – Duração 60 minutos – Estações do ano Inverno e Primavera.....	91

## LISTA DE TABELAS

Tabela 5.1 – Informações fisiográficas dos postos pluviográficos e estação de interesse .....	53
Tabela 5.1 – Parâmetros e características das distribuições <i>a priori</i> do limite superior .....	56
Tabela 5.2 – Características das distribuições <i>a posteriori</i> dos parâmetros do modelo LN4... 59	
Tabela 5.3 – Precipitação média anual e variação em relação ao observado para cada limiar 60	
Tabela 5.5 – Coeficientes do modelo de regressão para duração de 60 minutos .....	67
Tabela 5.6 – Resumo das probabilidades de similaridade encontradas.....	68
Tabela 5.7 – Probabilidades de similaridade por estação do ano – Duração de 60 min.....	70
Tabela 5.8 – Resultados das simulações de calibração – Parâmetros (S) e Limite – Duração 60 minutos .....	72
Tabela 5.9 – Resultados das simulações de calibração – Parâmetro (k) – Duração 60 minutos .....	73
Tabela 5.10 – Resultados das simulações de calibração – Duração 180 minutos .....	74
Tabela 5.11 – Resultados das simulações de calibração – Duração 360 minutos .....	74
Tabela 5.12 – Resultados das simulações de calibração – Duração 720 minutos .....	75
Tabela 5.13 – Resultados das simulações com 1.000 iterações para todas as durações.....	79
Tabela 5.14 – Comparação do modelo com uma Distribuição GEV – Duração de 60 min.....	83
Tabela 5.15 – Comparação do modelo com uma Distribuição GEV – Duração de 180 min...	84
Tabela 5.16 – Resultados das simulações – Diferenças entre modelos – Duração 60 minutos	84

## LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

AIC – Critério de informação de Akaike

ANA – Agência Nacional de Águas (ANA)

BIC – Critério de informação Bayesiano

CPRM – Companhia de Pesquisa de Recursos Minerais – Serviço Geológico do Brasil

CV – Coeficiente de variação

DJF – Dezembro, janeiro e fevereiro (estação do ano verão)

DP – Desvio padrão

EV4 – Distribuição de valores extremo do tipo IV

f, fx, FDP – Função densidade de probabilidade

F, Fx, FAP – Função acumulada de probabilidade

GAM – Distribuição Gama

GEV – Distribuição Generalizada de Valores Extremos

HPD – *Highest Probability Density* ou intervalo de credibilidade

IC – Intervalo de confiança

JJA – Junho, julho e agosto (estação do ano inverno)

k – número de vizinhos mais próximos

KDE – Estimativas de Densidade de Kernel

KNN – *K-nearest-neighbour* ou k-ésimo vizinho mais próximo

KNN-MOF – *K-nearest-neighbour – method of fragments* ou k-ésimo vizinho mais próximo – método dos fragmentos

KS – Kolmogorov-Smirnov

$k_m$  – fator de frequência

L – Função de verossimilhança

LN3 – Distribuição Log-normal de 3 parâmetros

LN4 – Distribuição Log-normal de 4 parâmetros

MAM – Março, abril e maio (estação do ano outono)

MCMC – Markov Chain Monte Carlo

MMM – Modelo Modificado de Markov

MPT – Modelo de matriz de probabilidade de transição

obs – valores observados

PMP – Precipitação Máxima Provável

RBLM – *Randomized Bartlett-Lewis mode*

S – número de estações vizinhas ou número de vizinhos

SNIRH - Sistema Nacional de Informação em Recursos Hídricos

SON - Setembro, outubro e novembro (estação do ano primavera)

TR – Tempo de retorno

TDF – *Transformed Distribution function*

u – Resposta binomial ( $u = 0$  ou  $u = 1$ )

$v_i$  - Características fisiográficas

WMO – Organização Meteorológica Mundial

$\alpha$  – Limite superior das precipitações máximas anuais

$\varepsilon$  – Limite inferior das precipitações máximas anuais

$\mu_y, \mu$  – Parâmetro de posição LN4

$\sigma_y, \sigma$  – Parâmetro de escala LN4

$\Phi$  – Distribuição normal padrão

$\rho$  – parâmetro de forma GAM

$\beta$  – parâmetro de escala GAM

# 1 INTRODUÇÃO

Séries de precipitação de alta resolução temporal são necessárias para a modelagem de eventos regulares e extremos em projetos de sistemas pluviais urbanos, simulação de vazões em pequenas bacias, estudos de balanços hídricos e modelagem de inundações de curta duração, por exemplo. Não obstante, séries temporais longas e contínuas de precipitação subdiária são, via de regra, difíceis de ser encontradas. Quando comparadas aos registros de precipitação diária, aquelas possuem, em geral, amostras de menor extensão, maior quantidade de dados faltantes e estações pluviográficas mais espaçadas geograficamente (WESTRA *et al.*, 2012).

Em virtude da disponibilidade limitada de registros pluviográficos, várias técnicas de desagregação de volumes precipitados, a partir de totais diários amostrados em pluviômetros, têm sido abordadas na literatura. Em sua ampla revisão sobre modelagem estocástica de precipitação, Sharma & Mehrotra (2010) indicam os métodos mais comumente utilizados para geração de chuva subdiária. Tais métodos compreendem variantes que utilizam processos de Poisson (processos de Neyman-Scott e de Bartlett-Lewis), modelos de desagregação baseados no princípio de invariância de escala (por exemplo, fractais e multi-fractais) e modelos não paramétricos baseados em técnicas de reamostragem (fragmentos).

Os modelos não paramétricos, segundo os referidos autores, são atrativos por serem flexíveis para desagregação de precipitação, conseguindo reproduzir as propriedades estatísticas da série observada e até, de certo modo, reproduzindo o comportamento dos máximos anuais para escalas temporais de menor duração, apesar da sua baixa capacidade de extrapolação e reprodução da variabilidade dos quantis. Além disso, o emprego de um modelo não paramétrico evita a criação de hipóteses tanto sobre a forma da distribuição do processo estocástico que origina as chuvas de duração subdiária, quanto sobre a relação entre as escalas temporais, o que reduz a incerteza do modelo (PUI *et al.*, 2012, SIVAKUMAR, 2017). Por fim, uma vez que certo nível de similaridade hidrológica entre as estações pluviográficas é estabelecido, a composição da informação regional para inferência é direta (WESTRA *et al.*, 2012; LI *et al.*, 2018), sendo útil para obtenção de informação de alta resolução temporal para a estação de interesse.

Dessa forma, por meio da obtenção da similaridade entre as estações pluviográficas, é possível realizar a geração de séries sintéticas de precipitação subdiária realizando a desagregação da



precipitação diária condicionada à altura de chuva e aos estados de ocorrência de precipitação do dia anterior e posterior no local de interesse, amostrando “fragmentos” subdiários obtidos das estações vizinhas (MEHROTRA *et al.*, 2012 e WESTRA *et al.*, 2012). Ao utilizar informações de outras estações pluviográficas, os modelos regionais conseguem aumentar a série de dados a ser reamostrada pelo modelo não paramétrico, melhorando a sua capacidade de reprodução das estatísticas das séries observadas. Entretanto, não conseguem contornar os problemas de reprodução da variabilidade dos quantis e baixa capacidade de extrapolação para quantis admitidos extremos. Com intuito de contornar essas dificuldades e obter estimativas mais confiáveis daqueles eventos de precipitação com reduzida probabilidade de superação, o modelo subdiário pode ser acoplado a um gerador estocástico de precipitação diária.

Assim como os subdiários, os geradores diários buscam reproduzir as propriedades estatísticas da série observada, só que, nesse caso, para a escala diária, utilizando métricas tais como: média, variância, número de dias secos e chuvosos e o comportamento dos extremos (WILKS & WILBY, 1999). Adicionalmente, a simulação de séries sintéticas por meio de geradores estocásticos possibilita quantificar as incertezas e avaliar o risco envolvido nas estimativas dos quantis raros e extremos de chuva. Os geradores estocásticos de precipitação diária têm sido amplamente pesquisados desde a década de 70 (BUISHAND, 1978; CHAPMAN 1994, 1997, 1998; SHARMA & LALL 1999; SRIKANTHAN & MCMAHON 1985 e 2001; WOOLHISER, 1992). De maneira geral, as abordagens diárias podem ser classificadas como modelos bipartidos, não paramétricos, matriz de probabilidade de transição (MPT) e híbridos (COSTA, 2015).

Destaca-se neste trabalho os modelos híbridos, cuja metodologia considera que os eventos regulares e extremos compreendem processos físicos distintos e, assim, as alturas de precipitação associadas a cada um desses processos são amostradas de populações diferentes (FURRER & KATZ, 2008). Esses modelos constituem uma alternativa eficiente para simulação simultânea de eventos regulares e extremos. Entretanto, os mesmos apresentam alguns problemas, como a definição do limiar entre os dois tipos de eventos e problemas de extrapolação, oriundos do comportamento da cauda superior da distribuição das precipitações diárias extremas.

Para contornar problemas de extrapolação, geralmente associados à utilização de distribuições de superior exponencial, alguns autores optaram por empregar distribuições de probabilidade com cauda superior potencial. Entretanto, seu uso proporciona geração de volumes de

precipitação fisicamente implausíveis com relativa frequência (CHEN & BRISSETTE, 2014). Uma solução encontrada para restringir as alturas de chuvas simuladas a valores fisicamente plausíveis foi a utilização de distribuições de probabilidade que apresentam em sua formulação um limite superior, como na distribuição Lognormal de 4 parâmetros (LN4), na distribuição de valores extremos do tipo IV (EV4) e na distribuição *Transformed Distribution function* (TDF) (COSTA, 2015).

No entanto, a definição de limites superiores de precipitação tem sido uma difícil tarefa na hidrologia. Botero (2006) sugeriu o uso de estimativas de Precipitação Máxima Provável (PMP) como um valor fixo para o limite superior. Contudo, estimativas dessa variável dependem das amostras disponíveis e das ferramentas de estimação (PAPALEXIOU & KOUTSOYIANNIS, 2006), o que lhes confere uma natureza “quase-determinística” (FERNANDES *et al.*, 2010). Devido às incertezas de estimação da PMP, essa é passível de ser interpretada de modo probabilístico, introduzindo as incertezas de sua estimação na análise. Para acomodar tais incertezas no modelo distributivo das precipitações diárias, Costa *et al.* (2015) recorreram ao paradigma Bayesiano para inferir uma distribuição *a priori* para o limite superior das alturas de chuva diária com base em estimativas de PMP.

Já para contornar o problema do limiar entre eventos regulares e extremos apresentado nos modelos híbridos, uma alternativa é utilizar de diferentes limiares e comparar com os resultados obtidos por meio de funções pré-definidas, como funções-objetivo (FURRER e KATZ, 2008). Esse método é considerado subjetivo, porém contorna o problema de continuidade entre as distribuições.

Contornando os problemas identificados com os modelos híbridos, pode-se obter um modelo de geração de precipitação diária, que consiga simular apropriadamente tanto os eventos regulares quanto os mais extremos. Esse aspecto do modelo diário permite que as séries diárias a serem desagregadas pelo gerador subdiário tenham o mesmo comportamento. Assim, neste trabalho, caracterizou-se a necessidade de acoplar ao modelo de desagregação subdiária utilizado, que consiste em um método não paramétrico de reamostragem juntamente com uma abordagem por similaridade regional, a um modelo de geração estocástica diária.

## 2 OBJETIVOS

### 2.1 *Objetivo geral*

O objetivo geral desta pesquisa é desenvolver um gerador estocástico de precipitação para a escala subdiária que seja apto a simular de maneira apropriada tanto os eventos regulares quanto aqueles admitidos extremos.

### 2.2 *Objetivos específicos*

Pretende-se que outros objetivos intermediários sejam alcançados. São eles:

- Eliciar a distribuição *a priori* informativa para o limite superior da distribuição LN4 utilizando estimativas de PMP;
- Definir o limiar entre eventos regulares e extremos por funções objetivo, em escala temporal diária, a partir da calibração de um modelo híbrido de geração estocástica;
- Identificar características fisiográficas que definam a similaridade da distribuição conjunta de atributos notáveis de precipitações subdiárias e volumes de chuva diária em pares de postos pluviográficos;
- Avaliar a variabilidade dos quantis utilizando o gerador acoplado em contrapartida ao gerador unicamente subdiário; e
- Avaliar a reprodução de extremos frente a análise de frequência convencional.

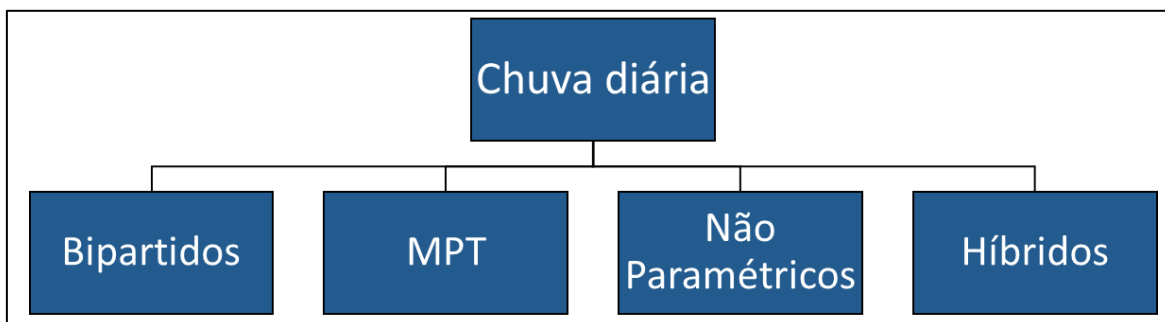
## 3 REVISÃO DA LITERATURA

### 3.1 Geração estocástica de precipitação diária

#### 3.1.1 Introdução

Para que os modelos de desagregação para a escala subdiária sejam eficientes, esses necessitam que as séries diárias a serem desagregadas sejam longas e contínuas, apresentando variabilidade suficiente para representar eventos de diferentes magnitudes. De modo a obter séries compatíveis com as premissas mencionadas, são amplamente utilizados os geradores estocásticos de precipitação diária, que buscam reproduzir propriedades estatísticas da série observada, tais como média, variância, número de dias secos e chuvosos e o comportamento dos extremos (WILKS & WILBY, 1999). Adicionalmente, a simulação de séries sintéticas por meio de geradores estocásticos possibilita quantificar as incertezas e avaliar o risco envolvido nas estimativas dos quantis raros e extremos de chuva.

Os geradores estocásticos de precipitação diária têm sido amplamente pesquisados e estudados (BUISHAND, 1978; CHAPMAN 1994, 1997, 1998; SHARMA & LALL 1999; SRIKANTHAN; MCMAHON 1985, 2001, 2003, 2004, 2005 a, b; WOOLHISER, 1992). Mais recentemente, Sharma & Mehrotra (2010) e Chen & Brissette (2014) apresentaram abrangentes revisões, as quais descrevem diferentes geradores de precipitação diária para regiões com condições climáticas diferentes. De maneira geral, tais modelos podem ser classificados assim como apresentado pela Figura 3.1: bipartidos, modelos não paramétricos, modelos de matriz de probabilidade de transição (MPT) e modelos híbridos. Tais categorias são apresentadas com detalhes nos tópicos seguintes.



**Figura 3.1** – Resumo dos principais modelos de geração de chuva diária

### 3.1.2 Modelos bipartidos

Os modelos bipartidos compreendem duas etapas de simulação: a primeira refere-se à modelagem da ocorrência de dias chuvosos; e a segunda trata da estimação das alturas de chuva nos dias em que ocorre precipitação. A sequência de dias chuvosos e secos é modelada por meio de processos estocásticos que consigam preservar as propriedades dos referidos estados de precipitação da série observada.

Com relação à ocorrência de precipitação, podem ser citadas técnicas de estimação como a renovação alternada, onde a modelagem da ocorrência de chuva é realizada por meio da alternância de intervalos de dias secos e chuvosos descritos por distribuições de probabilidade (WILBY *et al.*, 1998; WILKS, 1999), e cadeias de Markov, nas quais a modelagem se baseia nas probabilidades de transição entre estados de precipitação de dias sucessivos (GABRIEL & NEUMANN, 1962; BOUGHTON, 1999). Ambos conseguem descrever as estruturas de dependência dos eventos de chuva. Porém, para amostras de tamanho reduzido, o primeiro modelo está sujeito a maiores incertezas, oriundas do ajuste de distribuições de probabilidade para modelagem dos intervalos secos e chuvosos observados (WILKS & WILBY, 1999).

As Cadeias de Markov podem possuir níveis, sendo cada um deles associado ao número de dias consecutivos para o cálculo das probabilidades de transição. Por exemplo, Cadeias de Markov de primeiro nível estão associadas apenas com os estados do dia em questão e do dia anterior, já uma de segundo nível está associada aos estados do dia em questão e dos dois dias anteriores a esse. O uso de Cadeias de Markov de níveis superiores auxilia na obtenção de melhores características de variabilidade, porém necessita da estimação de um grande número parâmetros (WILKS & WILBY, 1999).

Para a segunda etapa, a estimação das alturas de chuva é realizada, via de regra, com auxílio de modelos paramétricos, ou seja, as inferências baseiam-se na escolha de uma única distribuição de probabilidade para geração dos volumes de chuvas. Chen & Brissette (2014), após analisar vários modelos discutidos na literatura, afirmam que as variantes mais utilizadas são a distribuição Exponencial, a distribuição Gama e a distribuição Lognormal. De acordo com Chowdhury *et al.* (2017), a distribuição Gama é uma das distribuições mais comumente utilizadas, já que possui dois parâmetros, que podem ser calculados pela média e pelo desvio padrão dos dias chuvosos.

Os modelos bipartidos usualmente se mostram aptos a reproduzir características da precipitação diária, em especial os dois primeiros momentos. Contudo, em virtude do decaimento exponencial da cauda superior das distribuições usualmente empregadas na modelagem, os mesmos são pouco efetivos para a reprodução dos eventos extremos e momentos de ordem superior. Nesse contexto, estudos recentes, como os de Papalexiou & Koutsoyiannis (2012 e 2016) e Papalexiou *et al.* (2013), têm demonstrado que os volumes de precipitação diária e os processos por ela governados são descritos de maneira mais efetiva por modelos distributivos dotados de cauda superior pesada.

### 3.1.3 Modelos não paramétricos e semi-paramétricos

Os modelos não paramétricos, por outro lado, simulam as alturas de chuva utilizando basicamente os dados observados, por meio de técnicas de reamostragem como o *bootstrap* e modelos estimação de densidades via kernel (KDE) (LALL & SHARMA, 1996).

Uma abordagem não paramétrica considerada promissora é o emprego do *k*-ésimo vizinho mais próximo (KNN) com a utilização da estimação de densidades via kernel proposta por Lall & Sharma (1996). Essa metodologia envolve um algoritmo de reamostragem com reposição baseado na técnica *bootstrap*. A escolha dos vizinhos mais próximos baseia-se em uma métrica para um vetor *D* de distância entre pontos no espaço. Em seguida, um dos vizinhos mais próximos é escolhido aleatoriamente para reamostragem e composição da série da estação de interesse.

Para a escala diária, outros autores também desenvolveram trabalhos promovendo ajustes e melhorias ao modelo inicialmente proposto (RAJAGOPALAN & LALL, 1999; BUIHAND & BRANDSMA, 2001; HARROLD *et al.*, 2003; SHARIF & BURN, 2007). Rajagopalan e Lall (1999) aplicaram a abordagem do KNN para a precipitação e outras cinco variáveis climáticas, demonstrando a habilidade do método em reproduzir as propriedades estatísticas amostrais. Buishand & Brandsma (2001) propuseram uma simulação *multi-site* da precipitação e temperatura diária para 25 estações. Harrold *et al.* (2003) focaram na questão das secas em seu trabalho, incorporando uma dependência de longo prazo ao modelo. Sharif & Burn (2007) desenvolveram um gerador climático capaz de gerar *inputs* para modelos hidrológicos em escalas diária e horária, utilizando do método KNN para reamostragem das séries históricas.

Os modelos não paramétricos também têm sido empregados em modelos de geração de séries de outras variáveis climáticas em escala local, como temperatura, temperatura do ponto de orvalho e velocidade do vento, dentre outras (GOYAL *et al.*, 2013). Para escalas espaciais mais amplas, essa classe de modelos necessita de informações climáticas de larga escala, previsões climáticas e projeções de mudanças climáticas, dentre outros (VERDIN *et al.*, 2015).

De maneira geral, modelos não paramétricos apresentam melhor eficiência que os demais para reprodução das características das alturas observadas, porém, via de regra, não são capazes de simular apropriadamente os eventos extremos, por causa de sua baixa capacidade de extrapolação (CHEN & BRISSETTE, 2014).

Os modelos desenvolvidos recentemente vêm buscando contornar os problemas de extrapolação e de correlação temporal apresentados pelos modelos não paramétricos. Variações aplicadas ao algoritmo KNN ou à junção com abordagens paramétricas, transformam os modelos paramétricos em modelos semi-paramétricos (SRIVASTAV & SIMONOVIC, 2015).

Nessa classe de modelos, podem ser mencionados estudos como o de Goyal *et al.* (2013), que utiliza do modelo do *k*-ésimo vizinho mais próximo juntamente com a abordagem de perturbação Kernel tal como sugerido por Salas & Lee (2010), e aquele de Mehrotra *et al.* (2015), que emprega Cadeias de Markov para geração da ocorrência de precipitação e o modelo de estimação de densidades via kernel (KDE) para geração dos volumes de precipitação.

Embora os modelos semi-paramétricos possuam, em geral, desempenho melhor que os não paramétricos, sua gama de parâmetros e algoritmos a serem implementados dificulta sua utilização, visto que se aumenta o tempo gasto para desenvolvimento dos modelos, tem-se maior custo computacional e necessita-se de séries mais longas para calibração dos parâmetros (CHEN & BRISSETTE, 2014).

#### **3.1.4 Modelos de matriz de probabilidade de transição - MPT**

Os modelos de matriz de probabilidade de transição (MPT) constituem uma variação dos modelos bipartidos, já que fazem uso de um conjunto maior de estados discretos, cada qual associado a um intervalo de alturas de chuvas. Tal expediente permite que a correlação entre os volumes precipitados em dias sucessivos seja preservada, de maneira mais efetiva, ao longo da simulação (WILKS & WILBY, 1999). Assim, os modelos MPT são uma Cadeia de Markov de primeiro grau com vários estados, sendo o estado 1 “seco” e os demais “chuvosos”. Os estados

chuvosos são modelados por distribuições uniformes ou lineares, com exceção do último estado, que originalmente é modelado por meio de uma distribuição exponencial ou por uma transformação Box-Cox (SHARMA & MEHROTRA, 2010). A transformação Box-Cox é utilizada para aproximar a distribuição dos dados de um modelo Gaussiano, nos casos em que os dados originais da amostra não seguem essa distribuição.

De acordo com Srikanthan *et al.* (2005), essa classe de modelos subestima, com relativa frequência, as variâncias nas escalas mensal e anual. De modo a contornar essa dificuldade, os autores utilizaram o ajuste desenvolvido por Boughton (1999), que consiste em um fator multiplicativo à altura diária estimada, permitindo igualar o desvio padrão anual estimado ao observado.

Em seu estudo, além de incluir o ajuste desenvolvido por Boughton (1999), Srikanthan *et al.* (2005) também utilizam uma simplificação no último estado da matriz. Ao invés de empregar a transformação Box-Cox do modelo original, eles utilizam da distribuição Gama para simular as alturas de chuva. Com os ajustes o modelo apresentou desempenho adequado, só não conseguindo representar bem a assimetria da precipitação mensal e o coeficiente autocorrelação anual. Contudo, os modelos MPT, assim como os modelos bipartidos e não paramétricos, apresentam limitações para a reprodução dos eventos extremos, devido principalmente, ao peso excessivo da cauda superior das distribuições usualmente empregadas no último estado da matriz de probabilidades de transição (e.g., a distribuição generalizada de Pareto) e ao número usualmente pequeno de observações de eventos de maior magnitude (CHEN & BRISSETTE, 2014).

### **3.1.5 Modelos híbridos**

Os modelos híbridos, por sua vez, assumem que eventos regulares e extremos compreendem processos físicos distintos e, assim, as alturas de precipitação associadas a cada um desses processos são amostradas de populações diferentes (FURRER & KATZ, 2008). Esses modelos constituem uma alternativa eficiente para simulação simultânea de eventos regulares e extremos.

Os modelos híbridos estritamente paramétricos constituem de uma boa opção para um gerador estocástico, ajustando-se bem aos dados observados e podendo estimar valores apropriados de extremos ao utilizar uma distribuição com cauda superior adequada. Entretanto, a determinação



de uma distribuição empírica para cada classe de eventos introduz incertezas à modelagem, sendo essas associadas à escolha da distribuição e à estimação dos parâmetros. Adicionalmente, a imposição de um limiar entre as duas distribuições do modelo híbrido pode fazer com que a densidade resultante seja descontínua nesse ponto de transição, o que inviabiliza a estimação de parâmetros pelo método de máxima verossimilhança e pode levar a instabilidades numéricas ao longo das simulações (FERRER e KATZ, 2008).

De modo a evitar a adição de incertezas com a escolha de uma distribuição e evitar problemas de descontinuidade, Costa (2015) propôs a utilização de um modelo não paramétrico para a simulação dos eventos regulares. Já para a simulação de eventos extremos, o autor manteve a abordagem paramétrica, visto que os modelos não paramétricos possuem baixa capacidade de extrapolação. Essa estrutura permite flexibilidade de ajuste, maior capacidade de reproduzir variâncias mensais e anuais e também possibilita a geração de eventos de grande magnitude com valores significativamente superiores aos das máximas alturas de precipitações observada (COSTA, 2015).

Como todos os modelos, os híbridos também apresentam algumas limitações e/ou dificuldades, como a definição do limiar entre os dois tipos de eventos e a definição da distribuição dos eventos extremos. Em relação à definição do limiar, uma ferramenta comum para sua definição é a utilização de diferentes limiares e a comparação com os resultados obtidos por meio de funções pré-definidas, como funções-objetivo (FERRER & KATZ, 2008). Como exemplos de funções-objetivo podem ser listadas a precipitação máxima anual, que reflete influências de eventos extremos, e as médias mensais e anuais, que refletem influências na variância para maiores durações, dentre outras estatísticas (COSTA, 2015). Esse método é considerado subjetivo, porém eficiente para distinção entre eventos regulares e extremos.

Outros autores também propuseram métodos diversos, como Wilson & Toumi (2005), que definiram em seu estudo que os eventos extremos seriam as precipitações diárias com probabilidade de excedência menor que 5%. Outros autores buscaram evitar a definição de um limiar, propondo um modelo de mistura de distribuições (WILKS & WILBY, 1999; BOUGUILA *et al.*, 2006; VRAC & NAVEAU, 2007; HUNDECHA *et al.*, 2009; NAVEAU *et al.*, 2016). Os modelos de misturas de distribuições evitam a definição de um limiar, já que a estrutura do modelo se baseia em uma função de ponderação que confere maior peso ao modelo com cauda superior pesada à medida que os valores das alturas de precipitação aumentam.

Entretanto, tais modelos apresentam complexidade na estimação das proporções de mistura e problemas de continuidade entre as distribuições (BOUGUILA *et al.*, 2006).

Em relação à definição de uma distribuição para os eventos extremos, essa é uma tarefa importante e complexa, já que pode subestimar ou superestimar os valores simulados. Por exemplo, a escolha de uma distribuição de cauda superior exponencial, em geral, subestima os eventos simulados. Já a utilização de uma distribuição de probabilidade de cauda superior potencial frequentemente proporciona geração de volumes de precipitação fisicamente implausíveis com relativa frequência (CHEN & BRISSETTE, 2014, COSTA *et al.*, 2015).

Uma solução encontrada para o peso excessivo da cauda superior, gerando volumes de precipitação implausíveis, foi a utilização de distribuições de probabilidade que apresentam em sua formulação um limite superior explícito, tais como a distribuição Lognormal de 4 parâmetros (LN4), a distribuição de valores extremos do tipo IV (EV4) e a distribuição *Transformed Distribution function* (TDF) (COSTA, 2015).

Costa (2015) trabalhou com as três distribuições mencionadas, avaliando o desempenho e as características de cada uma. O autor descartou a utilização da TDF após detectar problemas numéricos e avaliou o desempenho do seu gerador utilizando as outras duas distribuições. Nessa avaliação, a distribuição LN4 forneceu ajustes mais coerentes com os registros observados, ao passo que a distribuição EV4 apresentou aderência reduzida mesmo a eventos de menor magnitude. Além disso, a LN4 é assimétrica à direita e pode se comportar como uma distribuição de cauda pesada quando o limite superior especificado é muito elevado ou, ao menos, de magnitude muito superior aos registros observados (COSTA & FERNANDES, 2017). Assim, com base nas conclusões desses autores sobre a distribuição LN4, mais informações sobre essa distribuição são apresentadas a seguir.

Slade (1936) propôs a distribuição LN4 por meio da seguinte transformação de uma variável aleatória  $X$ :

$$Y = \ln \left( \frac{X - \varepsilon}{\alpha - X} \right) \quad (3.1)$$

na qual  $\varepsilon \in \mathfrak{R}$  é o limite inferior de  $X$ ,  $\alpha \in \mathfrak{R}_+$  é o limite superior de  $X$  e  $Y \sim \text{NOR}(\mu_y, \sigma_y^2)$ .

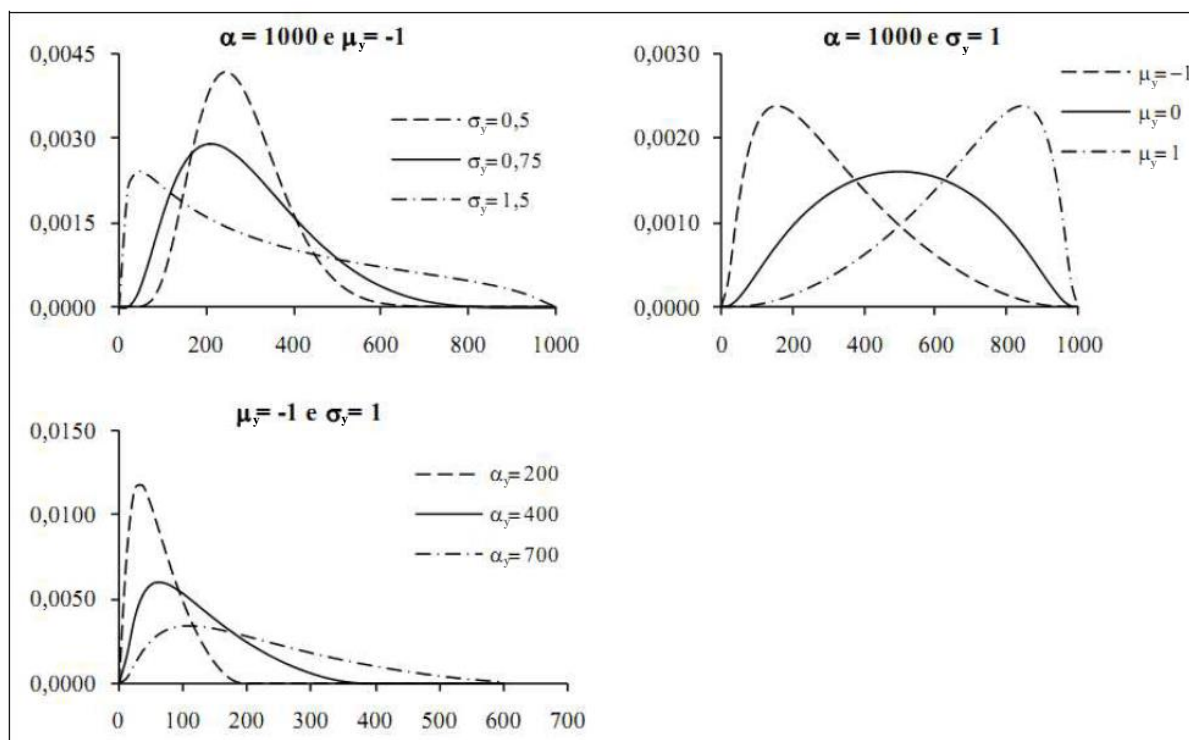
A função densidade de probabilidade da distribuição LN4 ( $\mu_y, \sigma_y, \alpha, \varepsilon$ ), com parâmetros de posição  $\mu_y \in \mathfrak{R}$ , parâmetro de escala  $\sigma_y \in \mathfrak{R}_+^*$ , limite superior  $\alpha \in \mathfrak{R}_+$  e limite inferior  $\varepsilon \in \mathfrak{R}$ , é dada pela Equação 3.2 e sua função de probabilidades acumulada é dada pela Equação (3.3), na qual  $\Phi$  denota a distribuição normal padrão.

$$f_X(x) = \frac{\alpha}{x(\alpha-x)\sigma_y\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma_y^2}\left[\ln\left(\frac{x-\varepsilon}{\alpha-x}\right) - \mu_y\right]^2\right\}, \varepsilon < x < \alpha \quad (3.2)$$

$$F_X(x) = \Phi\left[\frac{1}{\sigma_y}\ln\left(\frac{x-\varepsilon}{\alpha-x}\right) - \frac{\mu_y}{\sigma_y}\right], \varepsilon < x < \alpha \quad (3.3)$$

A primeira utilização da LN4 na hidrologia foi realizada por Takara & Loebis (1996) para a avaliação de eventos extremos de precipitação na Indonésia e no Japão. Os autores obtiveram melhores resultados com a LN4 do que com a distribuição log-normal de 3 parâmetros (LN3), a qual não é limitada superiormente.

No caso de precipitações, essas podem assumir somente valores não negativos. Desta forma, torna-se razoável admitir  $\varepsilon = 0$  para aplicações da LN4 com dados pluviométricos. Não é possível obter uma forma analítica para os momentos dessa distribuição. Assim, os mesmos devem ser obtidos por meio da solução numérica das equações de definição apresentadas anteriormente. A Figura 3.2 mostra a influência de cada parâmetro na forma da distribuição LN4. Nota-se pelos gráficos que os parâmetros  $\sigma_y$  e  $\alpha$  controlam o peso da cauda superior, sendo que, quanto maiores os valores desses parâmetros, mais pesada é a cauda. Já o parâmetro  $\mu_y$  controla a assimetria da distribuição. Para valores negativos de  $\mu_y$ , a assimetria da distribuição é positiva, enquanto que para valores positivos de  $\mu_y$ , a assimetria da distribuição é negativa; para  $\mu_y$  igual a zero a distribuição é simétrica.



**Figura 3.2** – Efeito de cada parâmetro na distribuição LN4

(Fonte: adaptado de Fernandes, 2009)

### 3.2 Limite superior de precipitações

A definição de limites superiores para precipitações e também para vazões têm sido uma difícil tarefa na hidrologia, devido a questões envolvendo a estimação desses limites, suas incertezas e até sobre a própria existência dos mesmos. Em relação a esse último aspecto, Botero & Francés (2010) argumentam que, considerando aspectos físicos, existe um limite máximo de escoamento gerado por uma precipitação máxima para determinada condição climática e características hidrológicas de uma bacia. Micovic *et al.* (2015), por sua vez, sugerem que existem três respostas possíveis para essa questão: 1) Existe sim um limite superior para precipitação; 2) Não existe limite, mas a taxa de variação da parte final da curva de precipitação-frequência é tão baixa que tende ao limite que é utilizado na engenharia; 3) Não existe limite e a taxa de variação da parte final da curva de precipitação-frequência indica que é possível ter precipitações maiores com probabilidade de ocorrência menores. Considerando o exposto, a questão da existência de um limite superior permanece controversa na literatura.

Assumindo sua existência, a Organização Meteorológica Mundial (*World Meteorological Organization* – WMO) propôs alguns métodos visando definir esse limite superior teórico para

precipitação. A WMO definiu a PMP (Precipitação Máxima Provável) como o limite teórico do volume de precipitação meteorologicamente possível em um determinado lugar, para uma dada duração e época do ano (WMO, 2009). A PMP pode ser estimada por métodos hidrometeorológicos e por métodos estatísticos. Os métodos hidrometeorológicos usualmente envolvem a maximização de tormentas extremas observadas no local ou transposição de tormentas. Já os métodos estatísticos utilizam de dados pluviométricos para estimação da PMP.

Em caso de indisponibilidade de dados hidrometeorológicos ou quando deseja-se obter uma estimativa rápida e preliminar de PMP, os métodos estatísticos constituem de uma alternativa simples. Entretanto, tal abordagem tende a gerar valores inferiores de PMP quando comparada com outros métodos (VMO, 2009). O método estatístico padrão da WMO é aquele desenvolvido por Hershfield (1961), com alterações no cálculo do fator de frequência. Estudos desenvolvidos pelo referido autor em 1965, mostraram que o fator de frequência é inversamente proporcional à magnitude do evento de chuva. Desse modo, o autor desenvolveu um ábaco para o cálculo do fator de frequência, com base na precipitação máxima anual e na duração do evento. A Figura 3.3 apresenta o ábaco, que encontra-se no manual da WMO. O procedimento de estimação da PMP é baseado na seguinte equação:

$$X_m = \bar{X}_n + k_m S_n \quad (3.4)$$

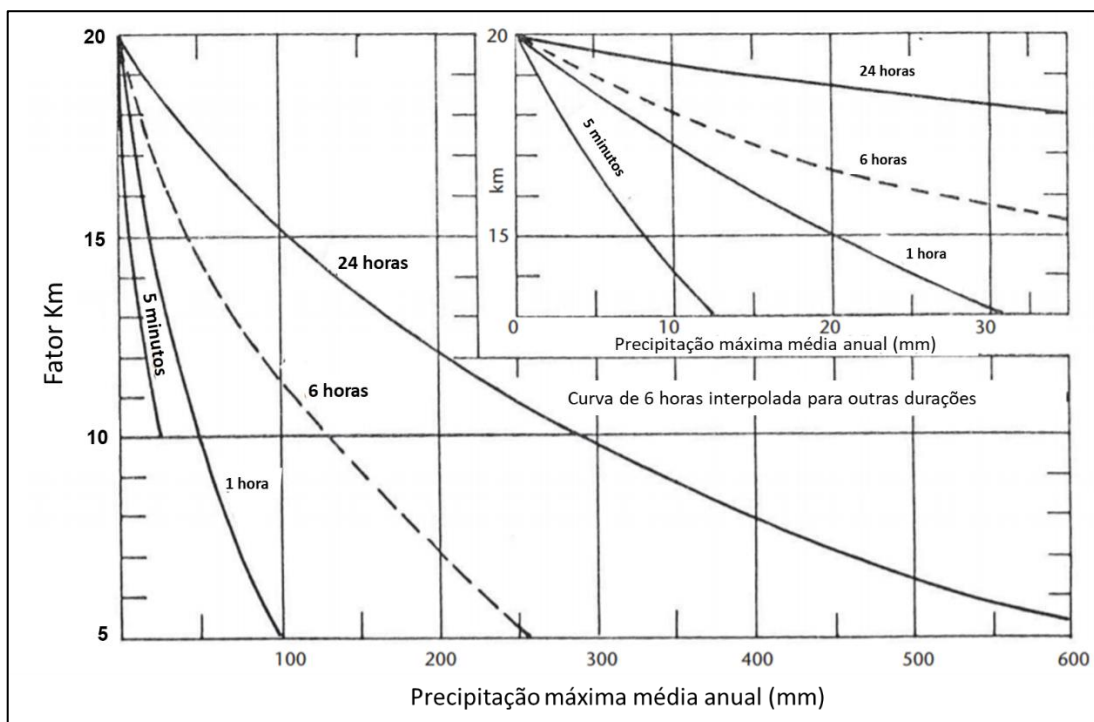
em que:

$X_m$  é a chuva máxima observada na bacia de interesse;

$\bar{X}_n$  é a média da série de precipitações máximas anuais;

$S_n$  é o desvio padrão da série de precipitações máximas anuais;

$k_m$  é o fator de frequência.



**Figura 3.3** – Estimativa do fator  $k_m$  para o método estatístico de estimação da PMP  
(Fonte: adaptado de WMO, 2009)

As estimativas de PMP obtidas com esse método são específicas para cada local. Portanto, as mesmas devem ser ajustadas para diferentes áreas de drenagem conforme as curvas altura-área-duração estabelecidas para a região de estudo (WMO, 2009). Além disso, alguns cuidados devem ser tomados para utilização desse método, como a presença de *outliers* na amostra, que pode levar à superestimação dos valores de média e desvio padrão máximos anuais, prejudicando o cálculo da PMP. Outra recomendação do manual da WMO (2009) é a utilização de séries longas, com no mínimo 20 anos de dados, para geração de estimativas mais confiáveis de PMP.

Koutsoyiannis (1999), utilizando a mesma base de dados de Hershfield (1961), realizou uma série de análises relacionadas à determinação do parâmetro  $k_m$ . O estudo demonstrou que não existem evidências acerca da existência de um limite superior para  $k_m$  e conseqüentemente para a PMP. O autor demonstrou ainda que a estimativa de PMP de Hershfield pode ser obtida a partir da distribuição GEV (Generalizada de Valores Extremos), com o parâmetro de forma dado por uma função linear da precipitação média anual máxima, e para período de retorno igual a 60.000 anos.

Devido à sua característica teórica e preliminar e considerando a disponibilidade limitada de séries históricas que apresentem eventos extremos, o método estatístico envolve um grau de subjetividade, resultando em incertezas a respeito dos valores pontuais obtidos, como discutido por Papalexidou & Koutsoyiannis (2006), Fernandes *et al.* (2010), Costa (2015) e Micovic *et al.* (2015).

Esses autores indicam que as estimativas de PMP são passíveis de ser interpretadas de modo probabilístico, introduzindo nas análises as incertezas de estimação. Micovic *et al.* (2015) sugerem, no mínimo, a utilização de uma faixa de variação para a PMP ao invés de uma estimativa pontual, para o caso de dimensionamento de estruturas que possuem alto custo e impacto associado, como as barragens e usinas nucleares. Já Costa (2015) recorreu ao paradigma Bayesiano para inferir uma distribuição *a priori* para o limite superior das alturas de chuva diária com base em estimativas de PMP, acomodando, assim, tais incertezas no modelo distributivo das precipitações diárias. A abordagem Bayesiana para inferência estatística é discutida a seguir.

### **3.3 Teoria Bayesiana para estimação de parâmetros**

Nas teorias frequentistas, os parâmetros de uma distribuição de probabilidades são quantidades fixas (não variáveis) e seus valores podem ser estimados, por exemplo, pela maximização da função de verossimilhança. Já na teoria Bayesiana, os parâmetros de uma distribuição probabilística podem ser interpretados como variáveis aleatórias, cujo conhecimento prévio é resumido por suas respectivas distribuições *a priori*. Distribuições *a priori* descrevem o estado de conhecimento sobre a quantidade aleatória antes da observação dos dados e pode ser representada por  $\pi(\theta/H)$ , sendo  $\theta$  o vetor de parâmetros e  $H$  a informação ou conhecimento prévio. À medida que se aumenta o nível de informação em relação ao parâmetro, espera-se que sua incerteza diminua e, no limite, ao menos em teoria, o total conhecimento do parâmetro implicaria em uma distribuição degenerada, na qual o parâmetro assume um valor único com probabilidade igual a 1 (COSTA, 2015).

Seja  $X$  uma variável aleatória relacionada a  $\theta$  e passível de ser amostrada. Supondo que o atual valor do parâmetro seja conhecido, a incerteza sobre a quantidade  $X$  é resumida pela função de verossimilhança  $f(X/\theta, H)$ , que fornece a probabilidade de ocorrência de cada amostra particular  $x$  de  $X$ . Com o auxílio do teorema de Bayes, a distribuição *a posteriori* do parâmetro

$\theta$ , que agrega o conhecimento atualizado sobre o mesmo e caracteriza completamente suas incertezas após a inferência, é dada por:

$$\pi(\theta/x, H) = \frac{f(x/\theta, H)\pi(\theta|H)}{f(x|H)} \quad (3.5)$$

na qual a função  $f(x/H)$  representa a distribuição preditiva *a priori* e é calculada por:

$$f(x|H) = \int_{\Theta} f(x/\theta, H)\pi(\theta|H)d\theta \quad (3.6)$$

A estimação de eventos extremos em engenharia e em outras disciplinas é realizada com base em informações incompletas e abstratas, existindo assim uma subjetividade intrínseca sobre o evento. Na teoria Bayesiana, essa subjetividade é analisada como o conhecimento prévio do especialista. Assim, a correta descrição da subjetividade, inerente ao processo natural de ocorrência de um evento extremo, depende da habilidade do especialista em selecionar, criticar, interpretar e julgar o conjunto de informações existentes sobre o evento (VICK, 2002).

Outro conceito introduzido pela teoria Bayesiana é o de intervalos de credibilidade, equivalentes aos intervalos de confiança (IC) dos métodos frequentistas. O intervalo de credibilidade para um parâmetro é construído com base na sua distribuição *a posteriori*, com a variância dessa distribuição fornecendo uma medida direta da incerteza do parâmetro. Portanto, esse leva em consideração apenas a amostra de fato observada, diferentemente do IC, que é baseado no princípio de repetição da amostra.

A maior dificuldade de utilização da abordagem Bayesiana é o cálculo da distribuição preditiva *a priori* dada na equação (3.6). Mais precisamente, para se fazer qualquer tipo de inferência sobre o modelo (momentos, quantis, intervalos de credibilidade), é necessário o cálculo do valor esperado de uma função  $h$  sobre a distribuição *a posteriori* dos parâmetros. Formalmente:

$$E[h(\theta)|x] = \frac{\int_{\Theta} h(\theta)f(x|\theta)\pi(\theta)d\theta}{\int_{\Theta} f(x|\theta)\pi(\theta)d\theta} = \int_{\Theta} h(\theta)\pi(\theta|x)d\theta \quad (3.7)$$

A solução da equação (3.7) envolve o cálculo de integrais multidimensionais complexas, as quais, via de regra, são impossíveis de serem obtidas por meios analíticos. A alternativa à integração analítica são os algoritmos de amostragem de Monte Carlo, mais especificamente aqueles baseados em cadeias de Markov (MCMC, do inglês *Markov Chain Monte Carlo*),



capazes de amostrar da distribuição *a posteriori* após a simulação de um grande número de realizações (GILKS *et al.*, 1996).

Com base em uma sequência de variáveis aleatórias  $\theta$  e uma função de interesse da população  $f(\theta)$ , a integração de Monte Carlo estima o valor esperado médio de  $f(\theta)$  utilizando amostras geradas  $\{\theta_i, i = 1, \dots, n\}$  da distribuição e aproximando  $E[f(\theta)]$  por:

$$E[f(\theta)] \approx 1/n \sum_{i=1}^n f(\theta_i) \quad (3.8)$$

O algoritmo MCMC gera amostras por meio de uma cadeia de Markov constituída ao longo tempo até que a distribuição estacionária seja atingida. As cadeias de Markov são processos estocásticos dotados de memória e que permitem a exploração do espaço paramétrico de maneira gradativa, o que, em geral, reduz as taxas de rejeição.

Para serem empregadas em algoritmos de amostragem, as cadeias de Markov devem ser:

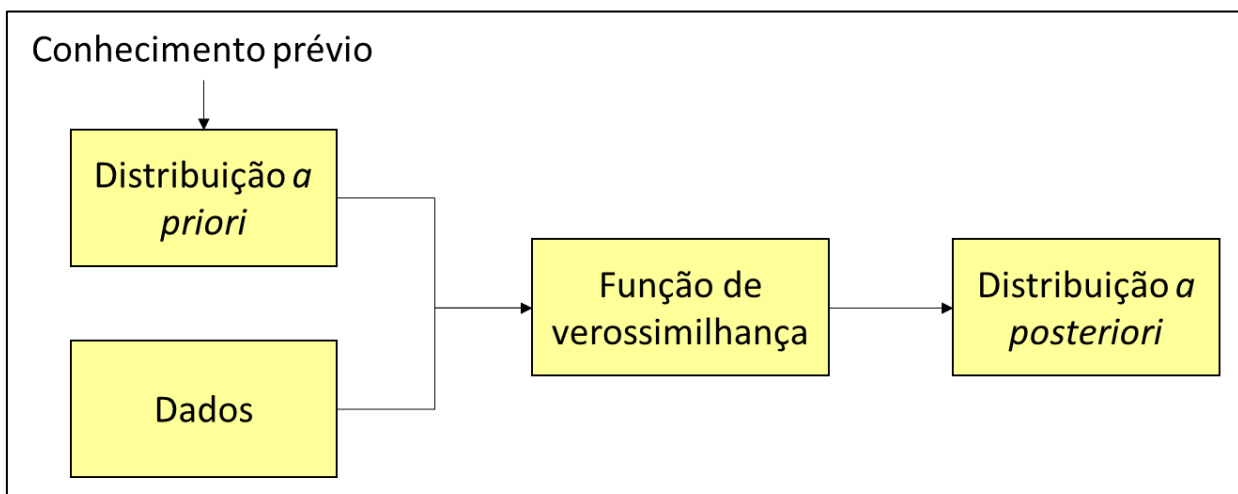
- Irredutíveis, o que significa que, independentemente de seu estado inicial, a cadeia é capaz de atingir qualquer outro estado em um número finito de iterações com probabilidade maior que zero;
- Aperiódicas, o que significa que a cadeia não oscila entre um conjunto de estados em movimentos regulares; e
- Recorrente, o que significa que, para todos os estados, se o processo se inicia em  $i$ , ele retornará a  $i$  em um número finito de iterações.

Uma cadeia de Markov com as características anteriores é denominada ergódica. A ideia básica de todos os algoritmos de amostragem é obter uma amostra da distribuição *a posteriori* construindo uma cadeia de Markov ergódica com as seguintes propriedades:

- A cadeia deve ter o mesmo número de estados de  $\theta$ ;
- A cadeia deve ser de fácil simulação; e
- A distribuição de equilíbrio deve ser  $\pi(\theta|x)$ .

No enfoque Bayesiano, a maneira usual de constituir cadeias é por meio de algoritmos como o amostrador de Gibbs e o algoritmo de Metropolis-Hastings (GILKS *et al.*, 1996). Esse último é construído por meio de uma distribuição de referência, da qual é fácil obter amostras de  $\theta$ , sendo sua escolha a principal questão para eficiência do algoritmo, pois quanto maior a semelhança entre a distribuição de referência e a distribuição alvo, mais rapidamente a cadeia alcança o equilíbrio (COSTA, 2015). Já o amostrador de Gibbs é um caso especial do algoritmo de Metropolis-Hastings, que permite gerar uma amostra da distribuição *a posteriori* desde que as distribuições condicionais completas estejam disponíveis para amostragem. O amostrador de Gibbs constitui uma alternativa de simulação interessante do ponto de vista computacional caso a obtenção das distribuições condicionais completas seja analiticamente possível. Mais informações sobre os algoritmos do método MCMC podem ser vistos por meio dos trabalhos de Gilks *et al.* (1996) e Robert & Casella (2004).

Em resumo, a teoria Bayesiana pode ser explicada pela Figura 3.5, na qual o conhecimento prévio do analista é utilizado para eliciação da distribuição *a priori* e juntamente com os dados amostrais, têm suas incertezas resumidas pela função de verossimilhança, tendo como resultado final a distribuição *a posteriori*, agregando o conhecimento atualizado e caracterizando completamente suas incertezas após a inferência.



**Figura 3.4** – Resumo da teoria Bayesiana

Após contornar os problemas de definição do limiar entre eventos regulares e extremos e a questão do limite superior da distribuição de probabilidade dos eventos admitidos mais extremos, incluindo nesse processo a teoria Bayesiana para incorporação das incertezas, pode-se obter um modelo de geração de precipitação diária, que consiga simular apropriadamente

tanto os eventos regulares quanto os mais extremos. O próximo passo é a definição de um modelo subdiário para a desagregação das séries diárias de precipitação. Para auxiliar essa escolha, serão apresentados a seguir os principais modelos discutidos na literatura.

### **3.4 Geração estocástica de precipitação subdiária**

#### **3.4.1 Introdução**

Desde a década de 1950, geradores estocásticos de precipitação têm sido utilizados para lidar com os problemas de disponibilidade limitada de dados de precipitação. Nessa classe de modelos, as propriedades estatísticas, tal como a média, variância, assimetria e autocorrelação, são extraídas de uma amostra e utilizadas para geração de séries sintéticas de qualquer extensão, nas quais os quantis de precipitação são gerados aleatoriamente por meio de funções de probabilidade de não excedência (COSTA *et al.*, 2015).

Recentemente, tem-se dado mais enfoque às séries de precipitação de alta resolução temporal, ou seja, subdiária, devido à sua importância para aplicação em projetos, estudos e gestão de recursos hídricos. Com séries longas e contínuas de precipitação subdiária, é possível um melhor planejamento, e a definição de estratégias de emergência e reposta em caso de precipitações extremas e/ou intensas de curta duração que impactam os sistemas de drenagem urbana, pequenas bacias, plantações agrícolas, levando a prejuízos/danos sociais, ambientais e econômicos (SHARMA & MEHROTRA, 2010). Entretanto, amostras de qualidade em escalas temporais refinadas, ou seja, longas e contínuas, são, via de regra, difíceis de ser encontradas, devido à disponibilidade limitada de registros pluviográficos e ao fato de que medições desse tipo de dado hidrológico tem um custo alto de tempo e de recursos financeiros.

Visando contornar essa dificuldade, várias técnicas de geração estocástica de precipitação subdiárias foram desenvolvidas nas últimas décadas (CHEN & BRISSETTE, 2014), auxiliadas pelo avanço dos recursos computacionais. Os modelos propostos possuem diversas estruturas e formulações, baseando na geração estocástica por meio de simulações diretas ou por meio de técnicas de desagregação de volumes precipitados a partir de totais diários amostrados. Cabe aqui destacar as principais classes de modelos apresentados na literatura, desenvolvidos ao longo dos últimos 50 anos: (1) modelos baseados em processos de Poisson (processos de

Neyman-Scott e de Bartlett-Lewis); (2) modelos baseados em invariância de escala (multi-fractais); (3) modelos paramétricos e (4) modelos não paramétricos (Fragmentos).

Os modelos baseados nos processos de Neyman-Scott e de Bartlett-Lewis (RODRIGUEZ-ITURBE *et al.*, 1987; COWPERTWAIT, 1991; ONOF *et al.*, 1995; GLASBEY *et al.*, 1995; ONOF *et al.*, 2000; KOUTSOYANNIS & ONOF, 2001; DINIZ, 2003; WHEATER *et al.*, 2005; KACZMARSKA *et al.*, 2014, 2015; KIM *et al.*, 2014, 2017; KOSSIERIS *et al.*, 2016; RAMESH *et al.*, 2018) tentam reproduzir o fenômeno de precipitação por meio de pulsos retangulares agrupados, cuja ocorrência é descrita por um processo de Poisson. Esses são utilizados para simulação direta de precipitação, precisando ser acoplado a outro método para desagregação para escala mais finas que 1 hora (KOUTSOYIANNIS & ONOF, 2001).

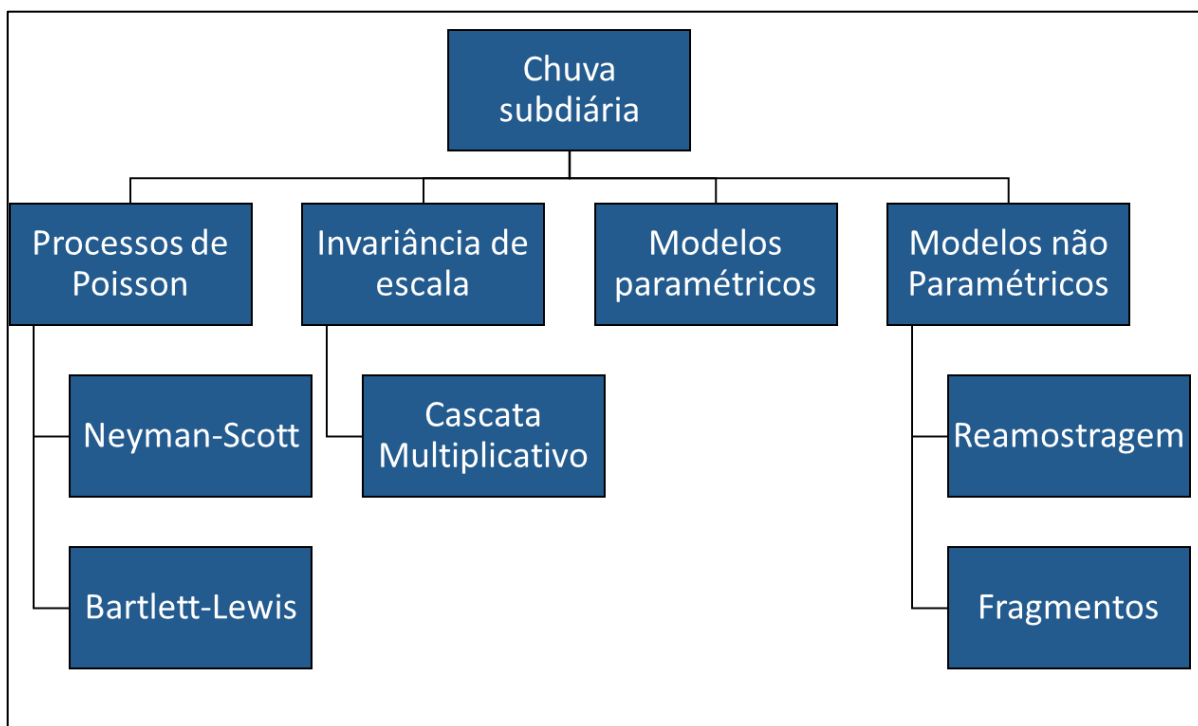
Os métodos baseados em invariância de escala e na teoria dos multi-fractais (Olsson *et al.*, 1993; OLSSON, 1998; VENEZIANO *et al.*, 2006; VENEZIANO & LANGOUSIS, 2010; DEIDDA *et al.*, 2006; KANG & RAMIREZ, 2010; RUPP *et al.*, 2012; MÜLLER & HABERLANDT, 2018; MCINTYRE & BÁRDOSSY, 2017) assumem um padrão de similaridade entre as escalas temporais, proporcionando uma distribuição uniforme da massa ou volume inicial, que é desagregada em estruturas menores sucessivas, para as quais uma quantidade da massa total é transferida.

Já os modelos paramétricos (HERSHENHORN & WOOLHISER, 1987; ECONOPOULY *et al.*, 1990; CONNOLLY *et al.*, 1998; DE MICHELE & SALVADORI, 2003 e 2005; VERNIEUWE *et al.*, 2015; PODUJE & HABERLANDT, 2017, 2018) atribuem funções de distribuições de probabilidade às características do evento chuvoso, como tempo de início, volume, intensidade e duração, tempo e volume do pico, tempo entre eventos de chuva, entre outros.

Por fim, os modelos não paramétricos (SVANIDZE, 1980; SRIKANTHAN & MCMAHON, 1985; WÓJCIK & BUIHAND, 2003; SHARMA *et al.*, 2006; WESTRA *et al.*, 2012; MEHROTRA *et al.*, 2012; LI *et al.*, 2018), baseiam-se em técnicas de reamostragem e do Método dos Fragmentos para desagregação da precipitação diária em subdiária.

A Figura 3.5 resume os principais modelos mencionados. Nos itens a seguir, as quatro classes de modelos são discutidas com mais detalhes, apresentando alguns dos trabalhos mencionados

acima, definições, conceitos, vantagens e desvantagens, limitações e outras informações pertinentes.



**Figura 3.5** – Resumo dos principais modelos de geração de chuva subdiária

### 3.4.2 Modelos baseados em Processos de Poisson

O Processo de Poisson possui características compatíveis com as exibidas pelo fenômeno da precipitação, sendo assim apropriado para utilização em modelagens hidrológicas. Os modelos baseados nesse processo conseguem preservar as características das precipitações em diferentes escalas de tempo e conseguem potencialmente representar as células de chuva, utilizando de premissas simples associadas ao processo físico (ONOF *et al.*, 2000).

Conforme Wheeler *et al.* (2005), desde o início da década de 1980 já eram desenvolvidos modelos utilizando Processos de Poisson, podendo destacar o trabalho de Rodriguez-Iturbe *et al.* (1987), que desenvolveram uma solução que permitiu a reprodução de importantes características da chuva em outras escalas temporais, como a altura e variância. Na década de 1990, vários modelos e trabalhos utilizando os modelos de Neyman-Scott e Bartlett-Lewis foram desenvolvidos, como listado por Sharma & Mehrotra (2010) em sua revisão da literatura: Cowpertwait (1991), Onof *et al.* (1995) e Glasbey *et al.* (1995). Modelos temporais-espaciais também foram desenvolvidos, utilizando dados de radar para identificar os processos de

formação dos grupos de células ou “*clusters*”. Entretanto, para que o modelo seja consistente com os valores de escalas temporais mais baixas (por exemplo, escala diária), o gerador deve ser ajustado para esses casos (KOUTSOYIANNIS & ONOF, 2001).

Mais recentemente, os trabalhos desenvolvidos têm visado aspectos como: a melhoria do desempenho para a desagregação de chuva em escalas temporais finas (KACZMARSKA *et al.*, 2014; KOSSIERIS *et al.*, 2016), regionalização dos modelos utilizando similaridade das propriedades estatísticas das estações (média, variância, autocorrelação e probabilidade de precipitação nula) (KIM *et al.*, 2017), diferentes tipos de pulsos, como o pulso de decaimento exponencial (RAMESH *et al.* 2018) e inclusão de não-estacionariedade (KACZMARSKA *et al.*, 2015). Destaca-se aqui, o trabalho de Kossieris *et al.* (2016), que utilizaram os ajustes incluídos por Koutsoyannis & Onof (2001) e a dependência entre a intensidade da chuva e duração da célula introduzida por Kaczmarska *et al.* (2014), para desagregação de chuvas em uma escala temporal mais fina (5 min).

Para representar apropriadamente as características da chuva, pode-se empregar modelos de diferentes níveis de complexidade, desde os mais simples, como os modelos *White Noise* e *Retangular Pulse*, àqueles mais complexos, como modelos de Neyman-Scott e Bartlett-Lewis. Os primeiros agregam a chuva em apenas um nível. Já os outros representam o fenômeno para uma ampla faixa de escalas temporais (DINIZ, 2003). Desse modo, os últimos conseguem representar as características estatísticas do fenômeno da chuva para diversas escalas temporais, atribuindo significado físico à estrutura de Pulso Retangular, ao associar a duração e intensidade de cada célula de chuva ao modelo.

Nos modelos de Neyman-Scott e Bartlett-Lewis a série de precipitação é considerada uma sequência de eventos de chuva, cada uma consistindo em um grupo de células de precipitação com uma intensidade e duração aleatória (MARAUN *et al.*, 2010). Basicamente são caracterizados por três processos estocásticos independentes, que são responsáveis por: (1) origem dos eventos, (2) número de células de chuva geradas a cada evento, e (3) origem das células. A diferença entre os dois modelos baseia-se no terceiro processo (origem das células), já que o tempo de chegada de cada célula de chuva é medida, no processo de Neyman-Scott, a partir da origem do evento, e no processo de Bartlett-Lewis, como o intervalo entre células sucessivas (SHARMA & MEHROTRA, 2010).

Geradores estocásticos baseados em processos de Poisson são considerados mais robustos e práticos que outros geradores, já que sua estrutura reflete bem os aspectos fundamentais do mecanismo de geração de chuva (KIM *et al.*, 2014). Entretanto, Koutsoyannis & Onof (2001) afirmam que, para esses modelos, balancear a matemática com a inclusão de parâmetros mais realísticos à modelagem é uma tarefa difícil, visto que modelos mais simples já envolvem uma gama considerável de parâmetros.

### 3.4.3 Modelos baseados em invariância de escala

Os modelos de cascata baseiam-se no comportamento invariante em escala do processo temporal de precipitação, ou seja, as propriedades estatísticas do processo de precipitação observadas em uma escala temporal maior podem ser utilizadas para estimação de propriedades estatísticas de uma escala mais refinada, visto que os dois são governados pela mesma relação de invariância de escala, permitindo, assim, o emprego dessa teoria para desagregação (OLSSON, 1998). O parâmetro de invariância de escala (fractal) é responsável pela ligação direta entre as propriedades estatísticas do processo em todas as escalas temporais. Como os processos atmosféricos são estruturas complexas, as propriedades estatísticas em diferentes escalas são relacionadas por meio de diferentes dimensões dependentes da intensidade, que podem ser interpretadas como um processo de cascata multiplicativo (teoria dos multi-fractais) (OLSSON *et al.*, 1993).

Os processos de cascata aleatórios multiplicativos estão associados com uma distribuição uniforme da massa ou volume inicial, que é decomposta em estruturas menores sucessivas, para as quais uma quantidade da massa total é transferida. Esses podem ser classificados como microcanônicos ou canônicos. A diferença é que o primeiro considera a conservação do volume de precipitação entre os níveis da cascata, já o segundo considera que apenas a média do volume é conservada (OLSSON, 1998).

Para considerar uma evolução temporal dos campos de precipitação é necessário um modelo mais completo, envolvendo “*space-time downscaling*”, ou seja, um modelo que seja capaz de obter informações hidroclimáticas para o local ou região da estação de interesse de modelos atmosféricos ou observações em uma escala espaço-temporal maior (RUPP *et al.*, 2012). Pode-se mencionar trabalhos considerando evolução temporal, como Deidda *et al.* (2006) e Kang & Ramirez (2010). Deidda *et al.* (2006) propuseram uma revisão da metodologia para caracterizar as propriedades de invariância de escala dos campos de precipitação no tempo e no espaço, e

aplicação do método para os dados de precipitação obtidos remotamente em duas campanhas nos oceanos e uma campanha sobre o continente. Já Kang & Ramirez (2010) propuseram uma abordagem mista, composta de um submodelo espaço-temporal estocástico, que preserva as características de dependência espacial e temporal entre escalas maiores e a escala de referência, e de um submodelo que preserva a similaridade estatística e a intermitência espacial entre escalas mais refinadas e a escala de referência.

Na literatura, alguns trabalhos, como o de Veneziano *et al.* (2006) e Veneziano & Langousis (2010), destacam-se pela extensa revisão, buscando apresentar os diferentes modelos, conceitos, tutoriais e trabalhos sobre os modelos utilizando fractais. Nos últimos anos, estudos envolvendo essa classe de modelos continuam sendo desenvolvidos, destacando-se os trabalhos de Müller & Haberlandt (2018) e de McIntyre & Bárdossy (2017) descritos mais detalhadamente a seguir.

Müller & Haberlandt (2018) analisaram três variações de modelo de cascata aleatório multiplicativo para desagregar a chuva diária, considerando uma consistência espacial na modelagem hidrológica de sistemas urbanos. Uma das conclusões desse trabalho é que sem a variabilidade obtida por reamostragem (adicionada em uma das variações do modelo), não se consegue obter valores apropriados de vazão nas modelagens.

McIntyre & Bárdossy (2017), visando melhorar a capacidade de estimação de extremos de um modelo de cascata aleatório discreto multiplicativo, empregaram duas técnicas: inclusão de estimativas de PMP (Precipitação Máxima Provável) e inclusão da dependência de volume nos parâmetros da FDP (Função Densidade de Probabilidade) dos Pesos. Os autores observaram que houve melhora na estimação de extremos no aspecto do volume de chuva. Entretanto, observaram uma piora no quesito frequência dos eventos extremos. Além disso, comparando os resultados do modelo com somente a dependência de volume e outro com a dependência e as estimativas de PMP, a utilização de somente da dependência de volume teve uma maior influência no resultado do que empregando as duas técnicas em conjunto.

Também foram desenvolvidos alguns estudos visando comparar modelos canônicos com modelos microcanônicos (LICZNAR *et al.*, 2011) ou comparar modelos multiplicativos de cascata com outros modelos (FERRARIS *et al.*, 2003). Cabe destacar que nesse último concluiu-se que o modelo multiplicativo de cascata apresentou um desempenho apropriado.;

Porém, após a otimização em relação as estatísticas amostrais (variância e curtose) e em relação



à dimensão fractal amostral, o mesmo não se comportou tão bem quanto o modelo utilizando processos de Poisson.

Os modelos de cascata são poderosas ferramentas para simular e descrever a complexidade atmosférica, incluindo a invariância de escala da precipitação no tempo e espaço (KANG & RAMÍREZ, 2010). Por ser mais simples e ter menos parâmetros, com a estrutura multi-fractal é possível obter as distribuições da intensidade e extremos da chuva (RUPP *et al.*, 2012). Embora sejam parcimoniosos (Sivakumar & Sharma, 2008), os modelos de cascata aleatórias frequentemente não conseguem reproduzir a variância observada, as alturas de precipitação extrema e a estrutura de dependência temporal da sequência completa de precipitação subdiária, o que muitas vezes limita sua aplicação em análises de risco de inundações e em dimensionamentos (DIEZ-SIERRA & del JESUS, 2019). Além disso, conforme salientado por Veneziano *et al.* (2006), não é tão fácil para o usuário aprender a teoria por trás desse tipo de modelo (fractais e multi-fractais), dificultando a análise dos processos atmosféricos.

#### **3.4.4 Modelos paramétricos**

A abordagem utilizada por modelos paramétricos faz uso de distribuições teóricas de probabilidade para ajustar os eventos de precipitação. Na literatura, versões iniciais de modelos paramétricos de desagregação podem ser vistas em Hershenhorn & Woolhiser (1987), Econopouly *et al.* (1990) e Connolly *et al.* (1998). Esses consistem de modelos simplificados quando comparados aos mais recentes, atribuindo funções de distribuições de probabilidade às características básicas do evento chuvoso, como tempo de início, volume, intensidade e duração. No entanto, a maioria das características utilizadas para descrever os eventos chuvosos possui relações de dependência entre si; por exemplo, um evento de longa duração é mais provável de estar associado a chuvas de baixa intensidade do que àqueles de grande intensidade (VERNIEUWE *et al.*, 2015). Desse modo, além de associar uma função distribuição de probabilidade a cada variável característica dos eventos de precipitação, deve-se considerar suas distribuições conjuntas.

Como as funções de distribuição de probabilidade marginais dessas variáveis não apresentam normalmente a mesma forma paramétrica, a utilização de funções de cópulas, capazes de modelar a dependência de variáveis aleatórias independentemente das suas distribuições marginais, aparece como alternativa (PODUJE & HABERLANDT, 2017). A teoria envolvendo funções de cópula foi desenvolvida por Sklar em 1959. Entretanto, segundo Vernieuwe *et al.*

(2015), apenas na década de 2000 é que foram realizadas as primeiras aplicações das funções de cópulas na hidrologia por De Michele & Salvadori (2003 e 2005). Conforme Poduje & Haberlandt (2018), a inclusão das funções de cópulas permitiu que os modelos desenvolvidos tivessem maior flexibilidade. Os modelos mais recentes já apresentam cópulas com até quatro dimensões, possibilitando a representação da estrutura de dependência completa das variáveis que descrevem o fenômeno de precipitação.

O desempenho de modelos paramétricos depende unicamente de se modelar apropriadamente as características do evento. Entretanto, tentativas de melhor representar o fenômeno resultam em parametrização em excesso, tornando os modelos mais complexos e necessitando de uma maior capacidade computacional. Tal fato torna a utilização desses modelos mais difícil em aplicações reais (SHARMA & LALL, 1999; PODUJE & HABERLANDT, 2018).

### 3.4.5 Modelos não-paramétricos

Os modelos não-paramétricos, ao contrário daqueles previamente citados, evitam a criação de hipóteses sobre a estrutura do modelo, a qual é definida unicamente pelos dados observados. Embora os métodos não paramétricos ainda envolvam parâmetros, o número e sua natureza não são pré-definidos, tornando-os flexíveis (SIVAKUMAR, 2017) e facilitando, desse modo, a modelagem e a aplicação dessa variante de modelagem por não envolver a complexidade da estimação de parâmetros.

Um dos métodos não paramétricos pioneiros é o “Método dos Fragmentos”, que foi inicialmente proposto por Svanidze em 1961 para geração estocástica de vazões (SVANIDZE, 1980) e tem sido amplamente utilizado na literatura tanto para desagregação de vazão quanto precipitação (LI *et al.*, 2018). O Método dos Fragmentos consiste na reamostragem de um vetor de fragmentos de precipitação subdiária, que correspondem ao valor percentual da precipitação diária, para a duração escolhida.

Lall & Sharma (1996) utilizaram de um estimador de densidade de kernel para obter os  $K$  vizinhos mais próximos *K-nearest-neighbour* - (KNN), e desenvolveram um algoritmo de reamostragem com reposição baseado na técnica *bootstrap*. A escolha dos vizinhos mais próximos baseia-se em uma métrica para um vetor  $D$  de distância entre pontos no espaço. Em

seguida, um dos vizinhos mais próximos é escolhido aleatoriamente para reamostragem e composição da série da estação de interesse.

Após o seu desenvolvimento, o algoritmo mencionado vem sendo aplicado por diversos autores com ajustes e modificações. Wójcik & Buishand (2003) utilizaram o algoritmo de reamostragem desenvolvido por Lall & Sharma (1996), mas com uma métrica diferente para o vetor D, para simulação de chuva com duração de 6 horas e também avaliaram a desagregação da precipitação diária para duração de 6 horas por meio do Método dos Fragmentos. Os autores concluíram que a abordagem de desagregação se comporta melhor que a simples reamostragem dos valores de precipitação com duração de 6 horas observados.

Sharma *et al.* (2006) fizeram a junção do Método dos Fragmentos e do algoritmo de reamostragem desenvolvido por Lall & Sharma (1996), acrescentando uma relação de dependência ao considerar os estados de chuva do dia anterior e posterior na desagregação de precipitação diária. O estudo de Sharma *et al.* (2006) também indicou algumas limitações dos modelos apresentados, como a dificuldade de se escolher o número de vizinhos mais próximos e o fato de que, por ser baseado em reamostragem dos dados observados, a ausência de séries longas e contínuas de precipitação de alta resolução temporal dificulta sua utilização. Para solucionar essa dificuldade, foram desenvolvidos modelos regionalizados, em vistas à ampliação dos dados para reamostragem, como apresentado pelos autores apresentados a seguir.

Visando expandir o trabalho desenvolvido por Sharma *et al.* (2006), Westra *et al.* (2012) e Mehrotra *et al.* (2012) apresentaram métodos regionalizados para a geração de séries sintéticas de precipitação subdiária em locais com e sem registros pluviométricos. O trabalho desses autores é um dos primeiros a acoplar uma abordagem regionalizada ao Método dos Fragmentos. No estudo de Westra *et al.* (2012), foram utilizadas 232 estações pluviográficas e no de Mehrotra *et al.* (2012) 2708 estações pluviométricas na Austrália. Os modelos foram validados utilizando cinco estações com séries longas e contínuas e os resultados encontrados sugeriram que ambos os modelos se mostram aptos a preservar a distribuição de probabilidade dos extremos e os momentos das séries observadas.

O estudo de Westra *et al.* (2012) foca na desagregação de precipitação diária para as durações de 6, 12, 30, 60, 120, 180 e 360 minutos, condicionada à altura de chuva e aos estados de ocorrência de precipitação do dia anterior e posterior no local de interesse, amostrando

fragmentos subdiários obtidos das estações consideradas vizinhas à estação de interesse. A definição de similaridade entre a estação em questão e as demais baseia-se em aspectos hidrológicos da duração escolhida (fração de zeros, máxima intensidade e tempo da máxima intensidade) e em aspectos fisiográficos das estações (latitude, longitude, elevação, distância da estação para a costa).

Primeiramente, a similaridade é avaliada em função dos aspectos hidrológicos e da precipitação diária, por meio do teste de Kolmogorov-Smirnov (KS) de duas dimensões e duas amostras desenvolvido por Fasano & Franceschini (1987). As estações consideradas similares para cada um dos aspectos hidrológicos são analisadas utilizando regressão logística, baseada em aspectos fisiográficos dos postos, obtendo-se assim probabilidades de similaridade das estações para cada atributo hidrológico. No trabalho de Westra *et al.* (2012), o atributo fisiográfico de maior influência foi a latitude, visto que estações utilizadas encontram-se espalhadas, na maioria, pela costa leste da Austrália, existindo, assim, um impacto maior da latitude do que nos demais atributos fisiográficos (latitude, elevação e distância para a costa).

Para a aplicação do algoritmo de desagregação, são consideradas aquelas estações com maior similaridade com o posto de interesse, em função de todos os aspectos hidrológicos. Após calcular os fragmentos de todas as estações, procura-se, para cada dia chuvoso, nas estações vizinhas, alturas de chuva similares à da estação em questão, considerando uma janela móvel de 15 dias e contemplando apenas os dias que tenham os mesmos estados (chuva/seco) do dia anterior e posterior. As alturas similares são classificadas em ordem crescente pela variação absoluta na altura de precipitação entre a estação de interesse e a estação vizinha, atribuindo uma probabilidade a cada uma delas. Aquelas estações com melhor classificação possuem maior probabilidade de serem similares. Por fim, é sorteada, com base nas probabilidades previamente definidas, uma estação com altura diária similar. Da estação escolhida obtém-se, então, o vetor de fragmentos correspondente, que é multiplicado pela precipitação do dia em questão na estação de interesse, conseguindo, assim, a série de precipitação subdiária para aquele dia.

Já o artigo de Mehrotra *et al.* (2012) foca na geração de precipitação diária regionalizada, para o caso de localidades não monitoradas, utilizando o mesmo algoritmo para identificação de quais estações são similares à estação de interesse. Para geração de precipitação diária, os autores utilizam o Modelo Modificado de Markov (MMM) – Estimativas de Densidade de

Kernel (KDE), tal como discutido por Mehrotra & Sharma (2007). A cadeia de Markov é condicionada ao estado do dia anterior, agregando as ocorrências de chuva a cada dia. Já os volumes são estimados por meio do procedimento KDE, considerando o estado do dia anterior. Por fim os autores também empregam o algoritmo de desagregação apresentado anteriormente.

Pui *et al.* (2012) compararam o desempenho de três modelos para desagregação de chuva diária em subdiária, quais sejam, um modelo multiplicativo aleatório de cascata, um modelo de Bartlett-Lewis aleatório “*randomized Bartlett-Lewis mode – RBLM*” e o modelo proposto por Sharma *et al.* (2006), chamado de “*K-nearest-neighbour – method of fragments (KNN-MOF)*”. Os autores concluíram que o modelo KNN-MOF tem desempenho superior em termos de reprodução dos momentos das séries observadas, das características dos extremos e das distribuições empíricas dos intervalos secos (*dry spells*) e chuvosos (*wet spells*) dentro de um dia. Lu & Qin (2014), por sua vez, confrontaram um modelo computacional chamado HYETOS, baseado no modelo de Bartlett-Lewis, com o modelo KNN-MOF. Os resultados da aplicação dos dois modelos também mostraram que o método KNN-MOF tem melhor desempenho para desagregação de chuva diária para horária na estação de interesse.

Mais recentemente, LI *et al.* (2018) desenvolveram três novas abordagens de desagregação, uma local, uma regionalizada e outra *multi-site*. Os três modelos utilizam do Método dos Fragmentos. Entretanto, os mesmos não utilizaram o algoritmo KNN e, sim classes de precipitação para agrupamento das alturas de precipitação, de modo que os fragmentos são sorteados da classe do dia em questão da estação de interesse. O modelo regionalizado é uma adaptação do modelo desenvolvido por Westra *et al.* (2012), já que os autores consideraram todas as estações como vizinhas da estação de interesse devido ao número limitado de estações. Já o modelo *multi-site* tenta manter a dependência espacial, desagregando todas as estações da região para o mesmo dia. Os três modelos foram aptos a reproduzir as estatísticas amostrais, indicando que as versões regionalizadas e *multi-site* apresentam melhor desempenho. Entretanto, os autores afirmam que os fragmentos reamostrados são limitados aos dados observados, indicando a baixa capacidade de extrapolação dos modelos.

Em virtude do exposto, os modelos não paramétricos possuem, para uma ampla gama de aplicações, desempenho relativamente superior ao comparado com outros modelos, além de serem fáceis de trabalhar, flexíveis e com custo computacional relativamente baixo. Portanto, o modelo a ser empregado nesta pesquisa consiste em um método não paramétrico de

reamostragem acoplado a uma abordagem de similaridade regional, na qual os “fragmentos” de precipitação subdiária são aleatoriamente amostrados de pluviógrafos nas proximidades, condicionados à altura de chuva diária no local de interesse. Entretanto, visando abranger a variabilidade decorrente da amostragem e se obter estimativas mais confiáveis daqueles eventos de precipitação com reduzida probabilidade de superação, será utilizado o gerador estocástico paramétrico de precipitação diária apresentado nos itens anteriores.

## 4 METODOLOGIA

A metodologia deste trabalho consiste em quatro etapas, as quais compreendem obtenção dos dados hidrológicos, o desenvolvimento do gerador estocástico de precipitação diária, o desenvolvimento do modelo de desagregação de chuva diária em subdiária e a validação e avaliação do desempenho dos modelos. Essas etapas são descritas a seguir.

### 4.1 *Etapa 1 – Dados Hidrológicos*

As informações necessárias para a realização deste trabalho foram obtidas junto aos órgãos responsáveis pela manutenção do Sistema Nacional de Informação em Recursos Hídricos (SNIRH), sob a responsabilidade da Agência Nacional de Águas (ANA) e operação do sistema pelo Serviço Geológico Brasileiro (CPRM). Em relação aos dados pluviográficos, os mesmos foram disponibilizados pela CPRM. Já os dados referentes às características fisiográficas, como latitude, longitude e elevação, foram obtidos da plataforma digital do SNIRH (<http://www.snirh.gov.br/hidroweb/publico/apresentacao.jsf>).

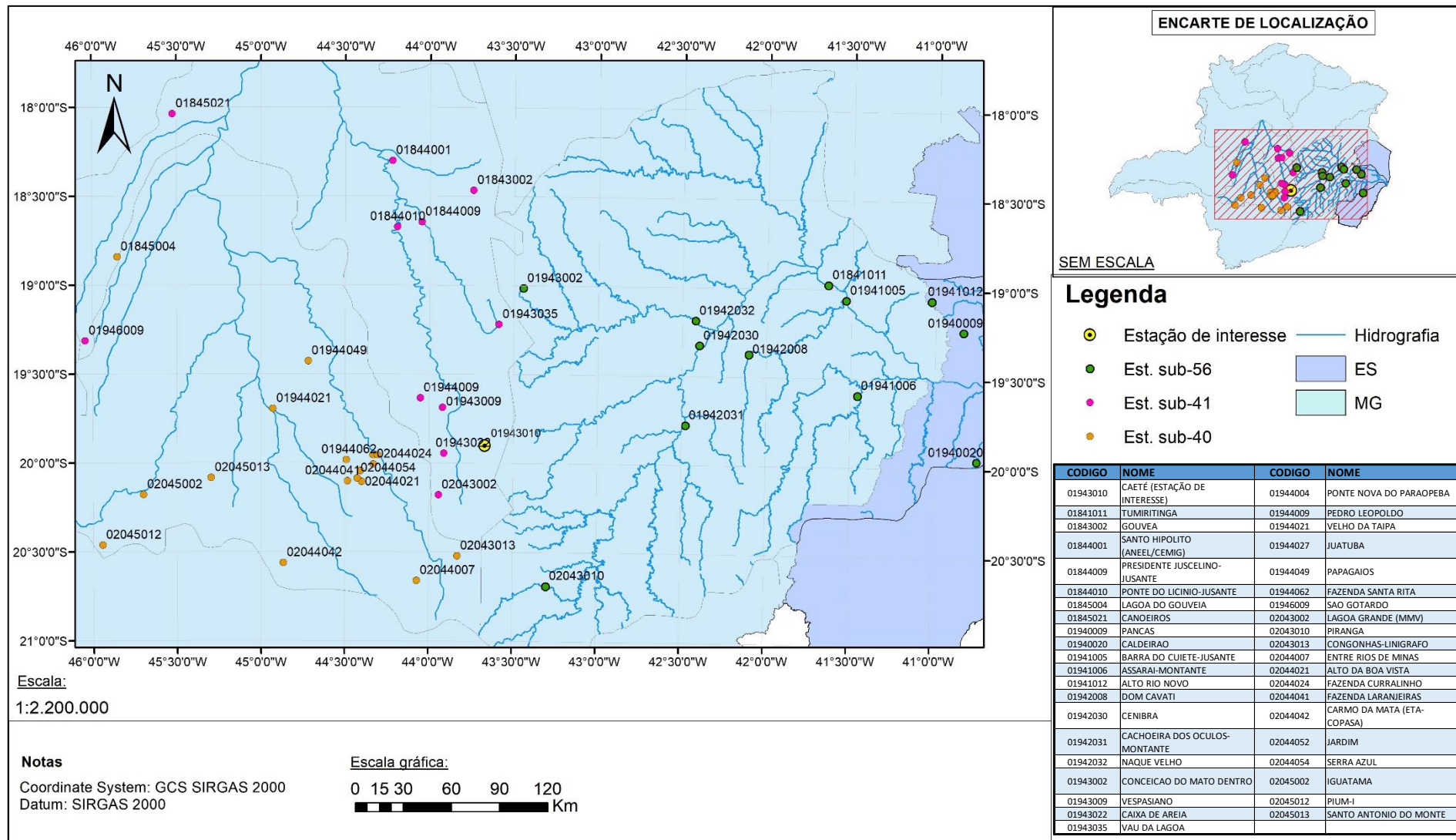
Em princípio, foram obtidos os dados de precipitação de 111 postos pluviográficos, assim como utilizado por Ferreira (2015). Entretanto, devido a problemas associados ao formato do arquivo do dado bruto fornecido, ao custo computacional para conversão e ao fato que, após convertidos para a extensão .txt, os arquivos ficam com mais de 1GB de tamanho para cada estação pluviográfica, o modelo subdiário ficou limitado em questão de tempo. Assim, para a realização do presente trabalho e desenvolvimento dos algoritmos, foram selecionados 40 dos 111 postos pluviográficos. Esses 40 são pertencentes às sub-bacias 40, 41 do rio das Velhas, inclusas na Bacia 4 - rio São Francisco e à sub-bacia 56 do rio Doce, inclusa na Bacia 5 – Atlântico, Trecho Leste (numeração conforme SNIRH). Esses postos foram utilizados no gerador subdiário para obtenção dos fragmentos utilizando a abordagem de similaridade regional. Os postos possuem séries curtas, com média de 6 anos de dados, registrados entre 2000 e 2008.

As falhas das estações utilizadas neste trabalho foram descartadas, já que o modelo proposto não retira amostras apenas de uma série e, sim, de várias estações. É necessário apenas garantir que a série diária a ser desagregada seja confiável e que existam informações suficientes dos demais postos pluviográficos para a amostragem de fragmentos. Assim, o impacto dessas falhas não é propagado para as séries desagregadas.

Para validação do modelo subdiário, tentou-se selecionar postos pluviográficos com séries de dados subdiários mais longas. Encontrou-se a estação de Caeté (Código 01943010), pertencente à bacia do rio São Francisco, sub-bacia 41 – rio das Velhas, que foi disponibilizada pela CPRM para realização do trabalho. Essa estação apresenta precipitações subdiárias observadas entre 1990 e 2017 (28 anos de dados). Devido à dificuldade de encontrar séries com essas características, apenas a estação de Caeté foi utilizada nesta etapa. A Figura 4.1 mostra as referidas sub-bacias no mapa e a locação das estações pluviográficas utilizadas. Essa figura indica também a localização da estação de interesse, a estação de Caeté (Código 01943010), utilizada para validação e avaliação do modelo.

Após a obtenção dos dados e descarte das falhas foi realizada uma avaliação preliminar dos dados obtidos, visando identificar possíveis anomalias. Os resultados dessa análise são apresentados no item 5.1, juntamente com mais informações sobre as estações pluviográficas.





**Figura 4.1** – Localização das estações pluviográficas utilizadas neste trabalho

## **4.2 Etapa 2 – Gerador estocástico de precipitação diária**

O procedimento do gerador diário descrito neste item é realizado de modo a abranger a variabilidade decorrente da amostragem e se obter estimativas mais confiáveis daqueles eventos de precipitação com reduzida probabilidade de superação. Assim, acoplou-se ao gerador estocástico de precipitação subdiária um gerador estocástico de precipitação diária. O gerador diário deve ser um modelo que consiga descrever tanto os eventos regulares e extremos. Como esses eventos compreendem processos físicos distintos, a escolha de um modelo que consiga abranger e simular ambos simultaneamente é essencial.

Os modelos híbridos conseguem realizar a simulação de alturas de precipitação para cada um dos processos ao obter amostras de populações diferentes. Nesse contexto, os modelos empregam distribuições distintas para cada tipo de evento, sendo uma para simulação de alturas de precipitação baixas à moderadas (regulares), via de regra associadas a eventos frontais, e uma segunda distribuição é utilizada para simular as chuvas extremas, oriundas de processos convectivos, mistos ou ainda aqueles resultantes da intensificação de convergências de massas de ar de diferentes características hidrometeorológicas.

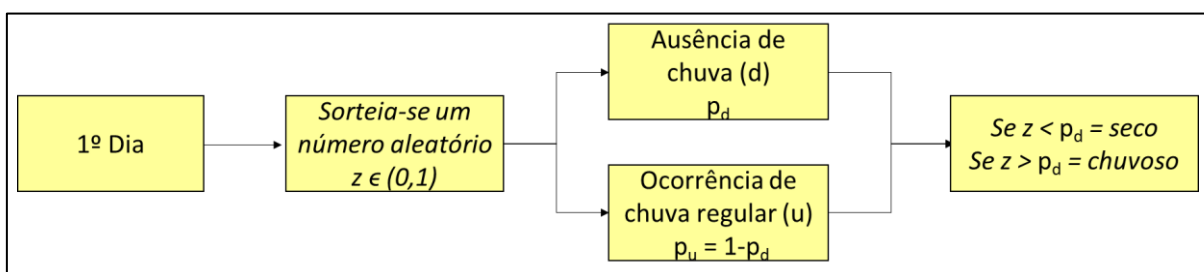
Seguindo a sugestão de Costa (2015), o gerador diário consiste de um modelo não paramétrico para a simulação dos eventos regulares e uma abordagem paramétrica para os eventos extremos. Para a ocorrência de chuva utiliza-se de uma matriz de probabilidade de transição de 3 estados, a saber: ausência de chuva, ocorrência de chuva moderada e ocorrência de chuva extrema, denotados, respectivamente, por (d), (u) e (e). Para a estimação das probabilidades de transição, foi empregado um modelo de cadeia de Markov de primeira ordem, no qual o estado de ocorrência de chuva no dia atual depende apenas do estado do dia precedente.

Por questões práticas, optou-se pelo modelo de primeira ordem, dada a dificuldade de estimação confiável dos parâmetros para modelos de ordem superior e, conforme discutido por outros autores, o desempenho de modelos de segunda ordem, baseados em índices como o critério de informação de Akaike (AIC) ou o critério de informação Bayesiano (BIC), não é, em geral, superior ao daqueles de primeira ordem, para uma ampla gama de bacias hidrográficas, com diferentes características climáticas e meteorológicas (Costa, 2015).

Para construção das matrizes de probabilidade de transição, utilizou-se a base diária de modo a suavizar a variação dessas probabilidades ao longo do ano, minimizando os efeitos das

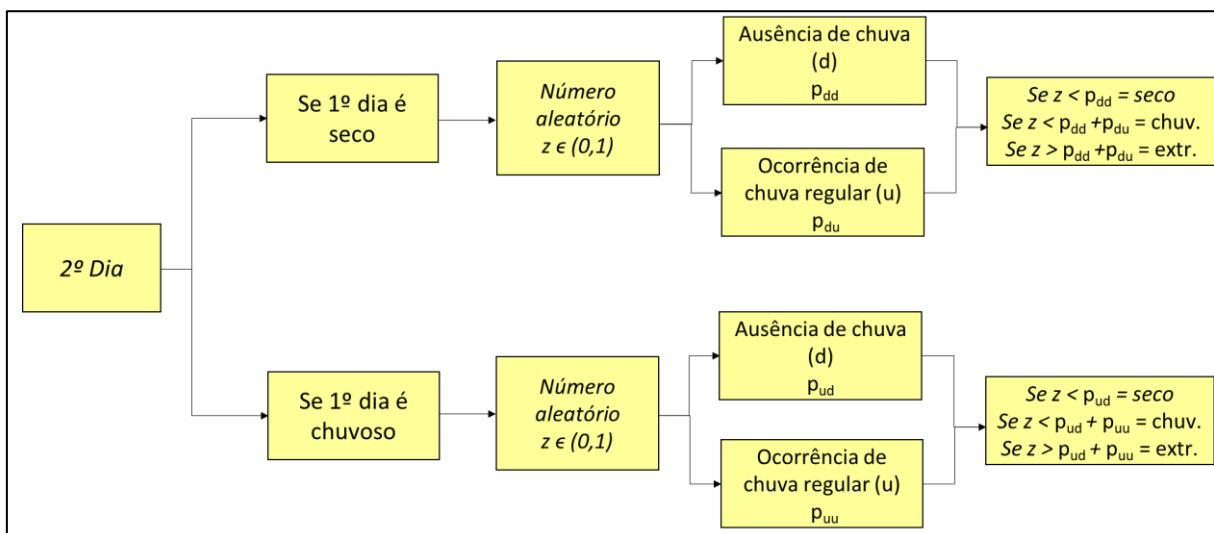
descontinuidades entre as mesmas na mudança entre intervalos discretos de grande duração, como quinzenas ou meses, durante a simulação. Assim, para cada dia do ano, com base na frequência histórica de transição e em um limiar estabelecido para diferenciar chuvas moderadas e extremas, são estimadas as probabilidades de transição.

Com auxílio de números aleatórios sorteados no intervalo (0,1) é realizada a seleção dos estados na matriz de probabilidade de transição, construindo, assim, a sequências de dias secos e chuvosos. A Figura 4.2 ilustra o procedimento para definição do estado do primeiro dia da simulação. Esse é definido com base em uma probabilidade de sucesso especificada, podendo essa ser obtida pela frequência histórica de dias secos. Nesse caso, se o número aleatório sorteado ( $z$ ) for inferior à probabilidade de sucesso ( $p_d$ ), o primeiro dia é seco. Caso contrário, o primeiro dia é chuvoso, e é considerada, por simplicidade, uma chuva convencional.



**Figura 4.2** – Procedimento para definição do estado do primeiro dia de simulação

Para determinar a ocorrência de chuva no segundo dia, utiliza-se o procedimento ilustrado pela Figura 4.3. Identifica-se o estado do primeiro dia e escolhe-se a linha correspondente na matriz de probabilidade de transição. Em seguida, sorteia-se um número aleatório e verifica-se sua posição na referida linha. Por exemplo, se o primeiro dia for seco, o segundo dia será seco se o número aleatório sorteado ( $z$ ) for menor que a probabilidade de transição de dia seco para seco ( $p_{dd}$ ). Se o número aleatório estiver entre  $p_{dd}$  e a soma de  $p_{dd}$  à probabilidade de transição de dia seco para chuvoso ( $p_{du}$ ), o segundo dia terá chuva convencional. Por fim, se o número aleatório for maior que  $p_{dd} + p_{du}$ , o segundo dia terá chuva extrema. O processo se repete para os dias seguintes.



**Figura 4.3** – Procedimento para definição do estado a partir do segundo dia de simulação

Após a definição da ocorrência das chuvas, a próxima etapa é a definição das alturas de chuva. As alturas de chuva baixas a moderadas são calculadas por meio de um procedimento de *bootstrap* convencional. O processo de reamostragem ocorre no interior de janelas móveis centradas no dia de interesse ao longo do período histórico. Assim, o procedimento é altamente influenciado pelo tamanho das janelas móveis de reamostragem, uma vez que uma janela de poucos dias pode não representar adequadamente a variabilidade das alturas de chuva para o dia em questão, enquanto uma janela de grande tamanho demandará esforço computacional significativamente maior.

Costa (2015) definiu o tamanho da janela por meio de funções objetivo, comparando a média e desvio padrão dos valores gerados para cada mês do ano com as referidas estatísticas dos registros observados. Tais índices foram aferidos para janelas de reamostragem de tamanho 8, 14 e 28 dias. Após as simulações, o autor observou que, para qualquer tamanho de janela, a média é adequadamente reproduzida. Todavia, para janela de 8 dias o modelo não apresenta ajuste para o desvio padrão, confirmando a hipótese que uma janela desse tamanho é incapaz de reproduzir a variabilidade das precipitações diárias verificada na amostra. Já as janelas de 14 e 28 dias apresentam comportamento similar, reproduzindo apropriadamente o desvio padrão. Entretanto, visto que a janela de 14 dias utiliza de um menor esforço computacional, a mesma foi escolhida no modelo. Frente ao exposto e por questões práticas manteve-se no modelo utilizado neste trabalho, a definição da janela de 14 dias.

Já para a modelagem das alturas de precipitação extremas foi empregada a distribuição Lognormal de 4 parâmetros (LN4). Essa distribuição foi escolhida devido ao seu melhor ajuste

aos dados observados em postos localizados na região de estudo, por possuir limite superior explícito, por ser assimétrica à direita e poder se comportar como uma distribuição de cauda pesada quando o limite superior especificado é muito elevado ou, ao menos, de magnitude muito superior aos registros observados, conforme apresentado por Costa (2015). Mais detalhes da distribuição LN4 podem ser vistos no item 3.1.5.

A hipótese de existência de um limite superior para a precipitação parece razoável dadas as condições climáticas de cada região. Entretanto, como apresentado em tópicos anteriores, esse é um assunto de discordância entre vários autores. Ainda que conclusões definitivas sobre a existência de um limite não tenham sido obtidas, autores como Li *et al.* (2012) e Chen & Brissette (2014) têm enfatizado que o uso de distribuições ilimitadas com caudas superiores demasiadamente pesadas leva à superestimação tanto da frequência quanto da magnitude dos eventos chuvosos extremos. Assim, de modo a evitar a geração de valores de precipitação improváveis na simulação de eventos extremos, modelos distributivos com um limite superior em sua formulação são utilizados.

Visando contornar a dificuldade de estimação e considerando a disponibilidade limitada de séries históricas que apresentem eventos extremos, pode-se utilizar da teoria Bayesiana para estimação do limite superior. Diferentemente da inferência estatística convencional, o método aqui empregado estima também as incertezas relacionadas ao parâmetro em questão (limite superior), por meio da especificação de uma distribuição *a priori* que agrega informações sobre os extremos com base nos conhecimentos dos processos físicos de formação das chuvas. Assim, a PMP (precipitação máxima provável), um estimador para o limite superior, foi incorporado ao modelo por meio de uma descrição probabilística completa, que acomoda as incertezas oriundas de sua estimação.

No contexto da presente pesquisa, o único parâmetro para o qual se tem informação disponível para a eliciação de uma distribuição *a priori* informativa é o limite superior. Para esse parâmetro, a distribuição *a priori* foi eliciada a partir: (1) de estimativas de PMP de caráter regional, fornecendo indícios acerca da variabilidade do referido parâmetro; e (2) da estimativa local de PMP associada a uma probabilidade de excedência, empregada como estimador do limite superior. Tais estimativas de PMP regional serão calculadas por meio de métodos estatísticos, que se baseiam nos dados pluviométricos existentes, levando em consideração as características do local de interesse (WMO, 2009). Esse método é uma alternativa à

indisponibilidade de estimativas obtidas por métodos hidrometeorológicos. Mais detalhes para construção da distribuição *a priori* para o limite superior são apresentados no item a seguir.

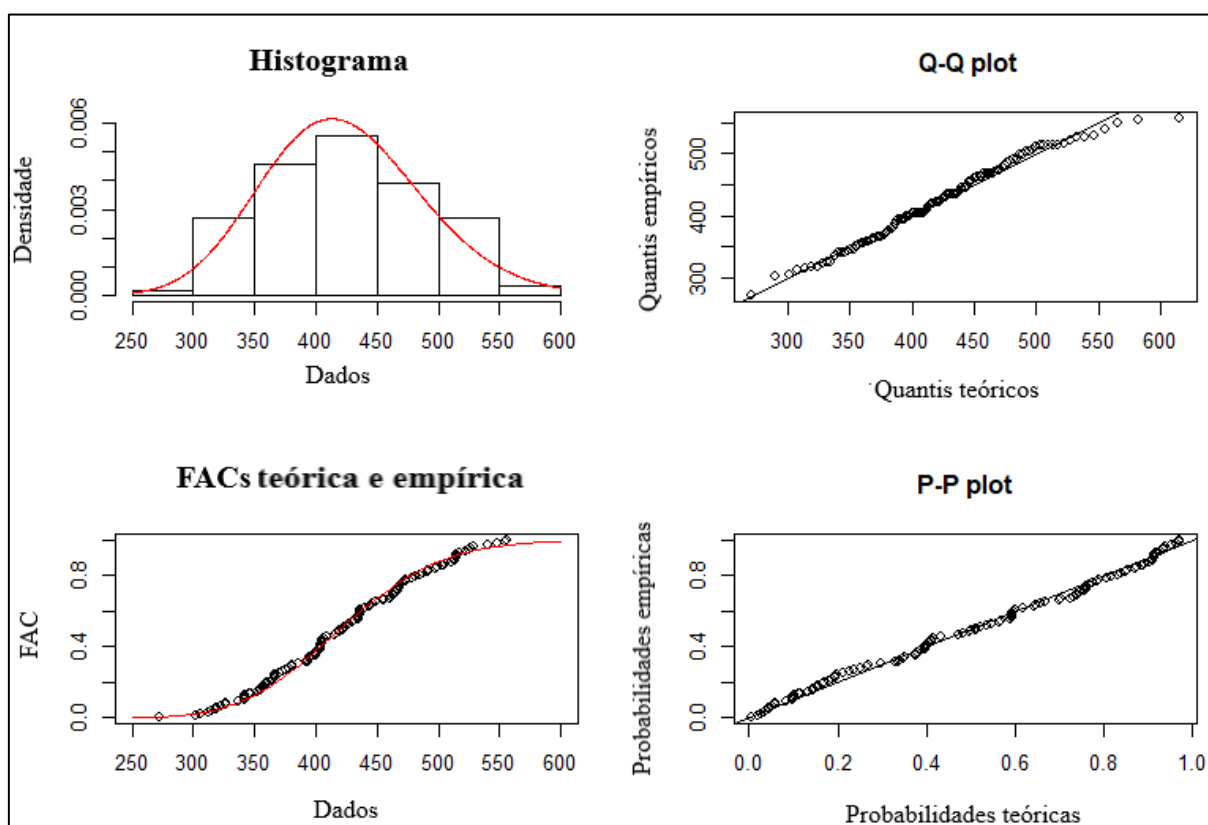
Outro aspecto importante nos modelos híbridos é a estimação do limiar entre os eventos regulares e extremos. Seu valor deve ser definido de maneira a se restringir a geração de valores de precipitação extremamente elevados em meses secos e, ao mesmo tempo, permitir a caracterização adequada dos eventos extremos nos meses chuvosos. Seguindo a sugestão de Costa (2015), foram utilizadas funções-objetivo para definição do limiar, sendo elas a precipitação máxima anual, que reflete influências de eventos extremos, e médias mensais, que refletem influências na variância para maiores durações.

#### **4.2.1 Construção das distribuições *a priori* dos parâmetros do modelo LN4**

Para construção das distribuições *a priori* dos parâmetros da distribuição LN4 o especialista baseia-se em seu conhecimento e em informações a respeito das características dos parâmetros. Esse procedimento atribui um grau de subjetividade à modelagem, constituindo assim de uma etapa importante no processo de inferência Bayesiana. Considerando os parâmetros da LN4, apenas para o limite superior ( $\alpha$ ) é que se tem informações disponíveis para construção de uma distribuição *a priori*. A PMP, como sugerido pelos autores mencionados no item 3.2, constitui-se de um indicador de limite superior. Entretanto, devido às incertezas relacionadas à sua estimação, é interessante utilizá-la de modo probabilístico, englobando as incertezas associadas. Para os demais parâmetros, a ausência de relação direta com o fenômeno físico faz com que sejam empregadas as chamadas distribuições *a priori* não informativas, caracterizadas por atribuir uma massa aproximadamente constante ao longo do espaço paramétrico.

Neste trabalho, escolheu-se a estação de Caeté (Código 01943010), pertencente à Bacia 4 - rio São Francisco, sub-bacia 41 – rio das Velhas (numeração conforme SNIRH) para validação dos geradores diário e subdiário. A estação apresenta dados de precipitação diária observados entre 1942 e 2018, totalizando 77 anos, e dados de precipitação subdiária observados entre 1990 e 2017. Porém, a mesma não apresenta estimativa de PMP hidrometeorológica. A situação ideal para avaliar a variabilidade da PMP e construir a distribuição *a priori* seria que uma grande amostra de estimativas de PMP, calculadas por métodos idênticos em diferentes épocas utilizando toda a informação hidrometeorológica, estivesse disponível.

Visto que tal situação não é possível, por motivos práticos, empregou-se o método estatístico utilizando um conjunto de estimativas de PMP de 118 estações pluviométricas em Minas Gerais, abrangendo uma diversidade de condições climáticas e hidrometeorológicas. O conjunto de estações e os valores das estimativas de PMP de 1 dia foram obtidos do trabalho de Costa (2015). Conforme discutido pelo autor, a distribuição Gama foi ajustada aos dados pela aderência às estimativas e à cauda superior exponencial, a qual atribui probabilidades reduzidas a valores implausíveis de precipitação, sendo indicada para modelar esse conjunto de dados. Visando obter uma visualização da distribuição dos dados, elaborou-se a Figura 4.4, a qual representa o ajuste da distribuição Gama pela linha contínua no histograma, juntamente com gráficos comparativos entre o empírico e o teórico, como o gráfico de quantis (Q-Q plot), funções acumuladas de probabilidade (FAPs) e de probabilidades (P-P plot).



**Figura 4.4** – Histograma de frequências das estimativas de PMP de 1 dia em Minas Gerais

As estimativas de PMP permitem avaliar, ao menos de maneira preliminar, a variabilidade do limite superior. De fato, se fossem obtidas as características climáticas e hidrometeorológicas do posto de interesse, provavelmente seria encontrado um histograma semelhante ao apresentado, possivelmente com menor variância. Baseando nessa suposição razoável, admitiu-se também a distribuição Gama para modelagem da distribuição *a priori* do limite superior.

Além disso, como a estação de interesse não possui informações de estimativa de PMP local por métodos hidrometeorológicos, buscou-se, em uma região próxima com características semelhantes, uma estimativa de PMP por esse método. Por fim, considerando a estimativa de PMP local como um estimador para  $\alpha$  e assumindo a distribuição Gama, é possível modelar a distribuição *a priori* do limite superior.

A adoção de uma distribuição de probabilidade ilimitada permite que sejam considerados incrementos de qualquer magnitude ao valor do referido parâmetro. Esse procedimento promove ao especialista a liberdade de não utilizar de um limite superior finito, que englobaria mais incertezas ao modelo, devido à falta de conhecimento prévio. Adicionalmente, uma distribuição com cauda exponencial, tal como a gama, atribui probabilidades bastante reduzidas àqueles valores de precipitação considerados implausíveis nas bacias em estudo, restringindo os valores amostrados a alturas mais prováveis (COSTA, 2015).

A função densidade de probabilidade de uma variável distribuída de acordo com o modelo Gama de dois parâmetros, a qual será denotada por  $X \sim GAM(\rho, \beta)$  com parâmetro de escala  $\beta \in R_+^*$  e parâmetro de forma  $\rho \in R_+$ , é dada pela seguinte equação:

$$f_x(x|\theta) = \frac{\beta^\rho}{\Gamma(\rho)} x^{\rho-1} \exp(-\beta x), x > 0 \quad (4.1)$$

Para estimação dos parâmetros da distribuição gama, empregou-se o coeficiente de variação regional (CV) e a frequência empírica de não-excedência da PMP local, já que não se possui uma série de PMP para a região de estudo.

O coeficiente de variação regional fornece a estimativa do parâmetro de forma. Pelo método dos momentos, tem-se:

$$\rho = \frac{1}{CV^2} \quad (4.2)$$

O parâmetro de escala  $\beta$  pode ser estimado admitindo-se uma probabilidade de não-excedência  $p$  para a estimativa local de PMP, ou seja, o parâmetro deve ser tal que  $P(\alpha \leq PMP/\rho, \beta) = p$ .

Associar uma probabilidade de não-excedência para a PMP, contudo, é uma tarefa complexa, mesmo quando estão disponíveis a melhor informação hidrometeorológica e ferramentas



adequadas para modelagem. Diante do exposto, a alternativa para se determinar uma probabilidade  $p$  de não-excedência da PMP local também pode ser fundamentada na análise regional. Nesse contexto, a ideia é se atribuir uma frequência de não-excedência empírica à estimativa local da PMP, com base em sua posição de plotagem no conjunto de 118 estimativas de PMP. As estimativas de PMP foram normalizadas pela chuva média anual, com objetivo de se extrair a influência climática da análise. Detalhes adicionais acerca das distribuições *a priori* serão apresentados no capítulo 5.

#### 4.2.2 Construção das distribuições *a posteriori* dos parâmetros do modelo LN4

A distribuição *a posteriori* dos parâmetros é proporcional ao produto da função de verossimilhança pelas distribuições *a priori*, ou seja:

$$p(\theta | x) \propto p(x | \theta)p(\alpha)p(\mu)p(\sigma) \quad (4.3)$$

onde  $p(x|\theta)$  é a função de verossimilhança,  $p(\alpha)$ ,  $p(\mu)$  e  $p(\sigma)$  são, respectivamente, as distribuições *a priori* para o limite superior, para o parâmetro de posição e para o parâmetro de escala.

A solução da Equação (4.3) envolve o cálculo de integrais multidimensionais complexas, as quais, via de regra, são impossíveis de serem obtidas por meios analíticos. A alternativa à integração analítica são os algoritmos MCMC, capazes de amostrar da distribuição *a posteriori* após a simulação de um grande número de realizações (GILKS *et al.*, 1996).

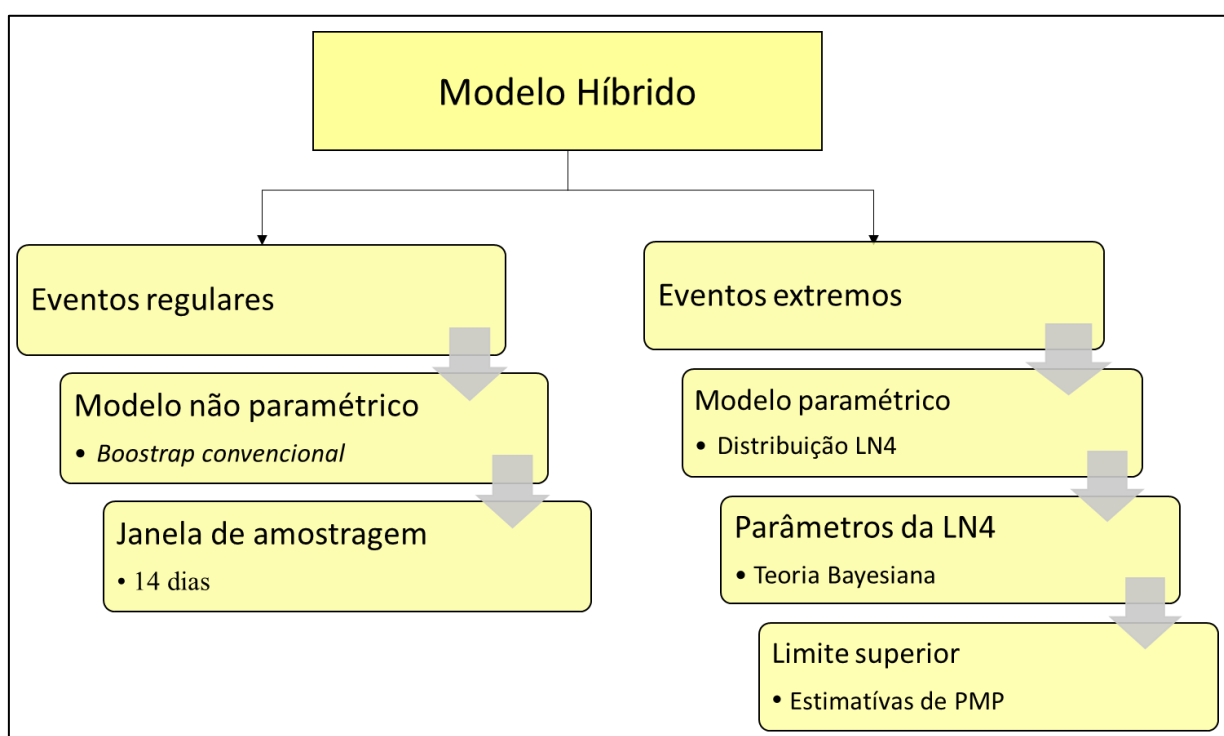
Desse modo, para estimação dos parâmetros, as simulações numéricas foram realizadas com auxílio do software OpenBUGS (LUNN *et al.*, 2009), que tem como uma de suas ferramentas o algoritmo MCMC. A utilização desse software possui as seguintes vantagens: (i) o programa permite o monitoramento de qualquer função dos parâmetros do modelo, (ii) fornece os valores amostrados de cada parâmetro monitorado, e (iii) fornece automaticamente resumos decorrentes da amostra obtida (média, desvio padrão e intervalo de confiança) (MATTOS & SILVA, 2002).

De posse das estimativas pontuais dos parâmetros da distribuição LN4, a função de quantis foi empregada para a simulação das alturas de chuva no algoritmo do gerador estocástico de precipitação diária.

### 4.2.3 Avaliação da eficiência do modelo diário

Foram geradas 1.000 séries com o mesmo tamanho da série histórica observada, avaliando-as por meio da comparação das estatísticas da série observada com estatísticas das séries geradas, como: (1) estatísticas diárias, como média, desvio padrão, coeficiente de assimetria, alturas máximas de precipitação, número médio de dias chuvosos para cada mês do ano; e (2) estatísticas mensais e anuais, como média e desvio padrão das precipitações médias mensais e anuais. As estatísticas mencionadas foram avaliadas considerando a média dos 1.000 valores, com exceção da altura máxima anual de precipitação diária, a qual foi avaliada para o conjunto da série de 1.000 valores.

Considerando o exposto, o gerador diário pode ser resumido pela Figura 4.5 e, após a confirmação do desempenho apropriado para o gerador diário, as 1.000 séries são empregadas no modelo do gerador de chuva subdiária para desagregação, conforme metodologia descrita no item 4.3 a seguir.



**Figura 4.5** – Resumo da metodologia do gerador diário

### 4.3 Etapa 3 – Geração dos dados em escala subdiária

A geração de chuva na escala subdiária tem como objetivo a geração de séries longas e contínuas, visto que os dados subdiários, via de regra, são mais escassos em termos de número de estações e de tamanho da série (menor que 10 anos de dados) (WESTRA *et al.*, 2012; PUI *et al.*, 2012; LI *et al.*, 2018). Esse aspecto pode ser observado pelo número de anos das séries diária e subdiária utilizadas neste trabalho e o baixo número de pluviógrafos quando comparado ao número de pluviômetros.

Conforme discutido por Pui *et al.* (2012), Lu & Qin (2014) e Li *et al.*, (2018), para geração de precipitação subdiária os modelos não paramétricos possuem desempenho relativamente superior ao de outros modelos, além de serem fáceis de trabalhar, flexíveis e com custo computacional razoável. Portanto, o modelo a ser empregado nesta pesquisa consiste em um método não paramétrico de reamostragem desenvolvido por Lall & Sharma (1996) acoplado à abordagem por similaridade regional proposta por Westra *et al.* (2012). A aplicação de um método de regionalização permite a utilização de dados de precipitação de estações vizinhas quando os dados da estação de interesse são muito curtos ou não se encontram disponíveis (WESTRA *et al.*, 2012).

A abordagem regional permite que os fragmentos de precipitação subdiária sejam aleatoriamente amostrados de pluviógrafos nas proximidades, condicionados à altura de chuva diária no local de interesse, assim como apresentado por Westra *et al.* (2012). Os chamados “fragmentos” correspondem ao valor adimensional de precipitação diária para a duração escolhida. Por exemplo, assumindo a duração de 1 hora e precipitação uniforme de 1 mm para cada hora, totaliza-se 24 mm de precipitação diária. Para obtenção do vetor de fragmentos basta realizar a divisão de cada valor de precipitação horária pelo total diário, assim tem-se um vetor de 24 fragmentos, cada um igual a  $1/24$ . Multiplicando-se a precipitação diária da estação de interesse por cada um dos fragmentos, obtém-se a série desagregada para a duração escolhida.

Para obtenção dos fragmentos, primeiro devem ser definidas as durações em que se deseja obter as séries subdiárias. Considerando o custo computacional do modelo proposto, que será discutido mais detalhadamente no item 5.3, e objetivando observar o comportamento em maiores durações, definiu-se pela utilização das durações de 60, 180, 360 e 720 minutos.

Após a obtenção dos fragmentos para cada uma das estações deste trabalho, deve-se determinar as características hidrológicas e fisiográficas que indiquem similaridade entre pares de estações pluviográficas. As características hidrológicas visam a identificação de correlações entre a precipitação diária e a subdiária, já as características fisiográficas buscam a expansão da similaridade para estações em que não se têm informações pluviográficas, ou seja, subdiárias.

#### **4.3.1 Definição da similaridade hidrológica**

As características hidrológicas devem ser definidas para cada uma das durações escolhidas (60, 180, 360 e 720 minutos), sendo que as utilizadas neste trabalho são:

- Intensidade máxima: para cada dia chuvoso, o intervalo com a máxima precipitação, que deve ser expressa de modo adimensional, ou seja, a razão entre a máxima precipitação e a precipitação total observada para aquele dia;
- Fração de zeros: para cada dia, a fração de intervalos sem precipitação; e
- Tempo da intensidade máxima: para cada dia chuvoso, o horário do dia em que ocorre a máxima precipitação.

Para os propósitos deste trabalho, assume-se que tais atributos incorporam um conteúdo de informação suficiente para a caracterização da chuva subdiária. Para identificar a similaridade hidrológica entre duas estações deve-se testar a hipótese de que a distribuição conjunta entre cada atributo hidrológico e as alturas de precipitação diária são estatisticamente similares. Assume-se que duas estações que possuem distribuições conjuntas similares possuirão também distribuições condicionais similares, já que o contrário pode não ser verdadeiro. Por exemplo, duas estações podem ter o regime de precipitação distribuídos em poucos intervalos (condicional similar), porém uma possui a maioria dos dias chuvosos com mais de 10 mm e outra possui a maioria dos chuvosos com precipitação diária menor que 10 mm (distribuições marginais diferentes). Assim, o foco não é na distribuição marginal, e sim na identificação dos pares de estações que condicionadas a uma mesma altura de precipitação diária possuem as mesmas características subdiárias.

Para testar a hipótese anteriormente descrita, será utilizado o teste de Kolmogorov-Smirnov (KS) de duas dimensões e duas amostras desenvolvido por Fasano & Franceschini (1987). A estatística do teste KS unidimensional é calculada obtendo-se o máximo das diferenças entre as

funções distribuição acumulada observada e empírica. Entretanto, para o teste bidimensional a função distribuição acumulada não é bem definida, devendo ser utilizadas as probabilidades integradas para os quatro quadrantes em torno de um ponto  $(x_i, y_i)$  em alguma dimensão  $x$  e  $y$  arbitrária. A estatística  $D$  do teste KS de duas dimensões é dada pela máxima diferença (compreendendo ambos os pontos dos dados e quadrantes) das probabilidades integradas. Tal quantidade é dada pela equação a seguir (PRESS & TEUKOLSKY, 1988).

$$P(D > observado) = Q_{KS} \left( \frac{D\sqrt{N}}{1 + \sqrt{1 - r^2 \left( 0,25 - \frac{0,75}{\sqrt{N}} \right)}} \right) \quad (4.4)$$

na qual,

$$N = \frac{N_1 N_2}{N_1 + N_2} \quad (4.5)$$

com  $N_1$  e  $N_2$  representando o tamanho das amostras 1 e 2, respectivamente. Para o cálculo da estatística do teste deve-se levar em conta esta equação:

$$Q_{KS}(x) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} e^{-2j^2 x^2} \quad (4.6)$$

Mais informações sobre a formulação das estatísticas do teste KS e detalhes sobre o mesmo podem ser vistos em Press & Teukolsky (1988). Para aplicação do teste será considerado um nível de significância de 5%. Como resultado do teste KS, tem-se uma resposta binomial,  $u$ , onde atribui-se 0 para não similares e 1 para os pares similares.

Visto que se tem como objetivo de expandir o método de similaridade para estações em que não há informações subsidiárias, além da similaridade hidrológica, deve-se verificar a similaridade fisiográfica.

### 4.3.2 Definição da similaridade fisiográfica

Os autores que desenvolveram o método de desagregação, a saber, Westra *et al.* (2012), utilizaram como características fisiográficas: a diferença absoluta entre elevações, latitudes,

longitudes, a distância até a costa marítima e a multiplicação das diferenças de latitude e longitude como características fisiográficas. Para o estudo de caso dos autores, a distância até a costa marítima tem algum impacto no modelo, já que foi utilizado um grande número de estações (232), com grande concentração na costa do país. Assim, os eventos chuvosos são influenciados pela proximidade ao mar. Entretanto, para o estudo de caso deste trabalho, com escala reduzida a três sub-bacias, para as quais quase todas as estações encontram-se distantes da região costeira, com apenas uma exceção, o emprego desse atributo não é interessante. Assim, as características fisiográficas empregadas foram: a diferença absoluta entre elevações, latitudes, longitudes e a multiplicação das diferenças de latitude e longitude, sendo essa última sendo uma alternativa para representação da distância entre as estações.

De modo a considerar tanto as características hidrológicas quanto fisiográficas, modela-se a resposta binomial ( $u = 1$ , similares; e  $u = 0$ , não similares) obtida das características hidrológicas com as quatro características fisiográficas ( $v_i$ ), por meio de um modelo de regressão logística:

$$P(u = 1) = \text{logit}(z) = \frac{e^z}{e^z + 1} \quad (4.7)$$

na qual,

$$z = \beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + \beta_4 v_4 \quad (4.8)$$

com  $\beta$  representando o vetor de coeficientes de regressão.

Dados os atributos fisiográficos, a Equação (4.7) fornece a probabilidade de que duas estações sejam similares, para cada atributo hidrológico separadamente.

Visando a definição de quais estações serão utilizadas na etapa de desagregação da chuva, considerou-se a média das probabilidades obtidas para cada atributo hidrológico. Aquelas estações com maior similaridade, considerando todos os atributos hidrológicos, são empregadas no algoritmo de desagregação. Vale ressaltar que, de modo a considerar as influências sazonais nas variações das relações entre precipitações diária e subdiária, separou-se o modelo para cada estação do ano. Assim, estações pluviográficas diferentes podem ser selecionadas para desagregação para cada estação do ano.

### 4.3.3 Algoritmo de desagregação da chuva diária em subdiária

Após a obtenção dos fragmentos e da identificação das estações que são consideradas similares à estação de interesse para cada estação do ano, aplicou-se o modelo de desagregação. O modelo proposto por Westra *et al.* (2012) consiste em uma variação do método dos fragmentos proposto por Lall & Sharma (1996). A primeira modificação se dá por meio da abordagem regionalizada, na qual os fragmentos de precipitação subdiária são aleatoriamente amostrados de pluviógrafos nas proximidades, condicionados à altura de chuva no local de interesse. A segunda modificação é que os fragmentos são amostrados considerando também as condições dos estados do dia anterior e dia seguinte.

O algoritmo de desagregação funciona da seguinte maneira:

1. Obtêm-se os dados de precipitação diária observados nas estações e calculam-se os fragmentos para cada estação, dia e duração, por meio da seguinte equação, a qual expressa os fragmentos como um percentual da precipitação diária:

$$fr_{i,m}^s = \frac{X_{i,m}^s}{\sum_m X_{i,m}^s} \quad (4.9)$$

na qual  $X_{i,m}^s$  representa a altura de chuva de uma estação  $s$ , em um dia  $i$  e em uma duração  $m$ .

2. Identificam-se as estações vizinhas (S) da estação de interesse para cada estação do ano por meio da abordagem regional apresentada nesta seção;
3. Para cada dia chuvoso, procuram-se nas estações vizinhas, ao longo de toda a série, alturas de chuva similares à da estação em questão. Considerou-se uma janela móvel de  $\pm 15$  dias para preservar as características sazonais dos pulsos de chuva, procurando em todos os anos apenas os dias que tenham os mesmos estados (chuva/seco) do dia anterior e posterior;
4. Contam-se os dias  $j = 1, \dots, n$ , dentro da janela móvel, que possuam os mesmos estados do dia anterior e posterior ao dia de interesse, em todas as séries observadas. Classificam-se então esses dias em ordem crescente, por meio da variação absoluta na altura de precipitação, ou seja, diferença de precipitação entre a estação de interesse e a estação vizinha. Usa-se a simbologia de parênteses para indicar os dias já ordenados (j);

5. Considerando um limite da variação absoluta de 5% e 10% (a ser definido de acordo com o desempenho do gerador), encontram-se os (k) vizinhos mais próximos, ou seja, aqueles dias que atendem o limite de variação, ordenados de maneira crescente;
6. Sorteia-se um número aleatório entre 0 e 1 e compara-se com a probabilidade calculada pela equação a seguir para decisão de qual dia (j) serão utilizados os fragmentos (LALL & SHARMA, 1996; MEHROTRA & SHARMA, 2006):

$$P(j) = \frac{1/(j)}{\sum_{i=1}^k 1/i} \quad (4.10)$$

8. Calcula-se a altura de chuva subdiária multiplicando o fragmento selecionado com o passo anterior pela altura de chuva do dia em questão simulada pelo gerador estocástico diário; e
9. Repetem-se os passos 3 ao 7 até que todos os dias de cada uma das 1.000 séries geradas pelo gerador diário sejam desagregadas.

O procedimento mencionado deve ser repetido para cada uma das durações escolhidas (60, 180, 360 e 720 minutos), obtendo-se, assim, as séries desagregadas para cada duração. É importante ressaltar que o algoritmo descrito depende da definição de alguns parâmetros, como o limite da variação absoluta, o número máximo de vizinhos (k), o número de estações vizinhas (S) à estação de interesse e o tamanho da janela móvel.

De acordo com Westra *et al.* (2012) o maior desafio desse procedimento é a definição do número de estações vizinhas (S), já que existe um grau de subjetividade associado a essa escolha. Um número grande de estações pode acarretar na seleção de fragmentos de estações estatisticamente diferentes da estação de interesse e um número baixo resulta em poucas estações para amostragem dos fragmentos.

Dessa forma, para o parâmetro número de estações vizinhas (S), fez-se a opção por avaliar (S) igual 5, 10 e 15, tendo-se em vista o número limitado de estações para a desagregação (40 estações) e o custo computacional. Tal estratégia teve fundamentação nos trabalhos de Westra *et al.* (2012), Pui *et al.* (2012) e Li *et al.* (2018) com a definição de 13, 4 e 6 estações vizinhas, respectivamente. No caso dos dois últimos, os autores só possuem esse número de estações para

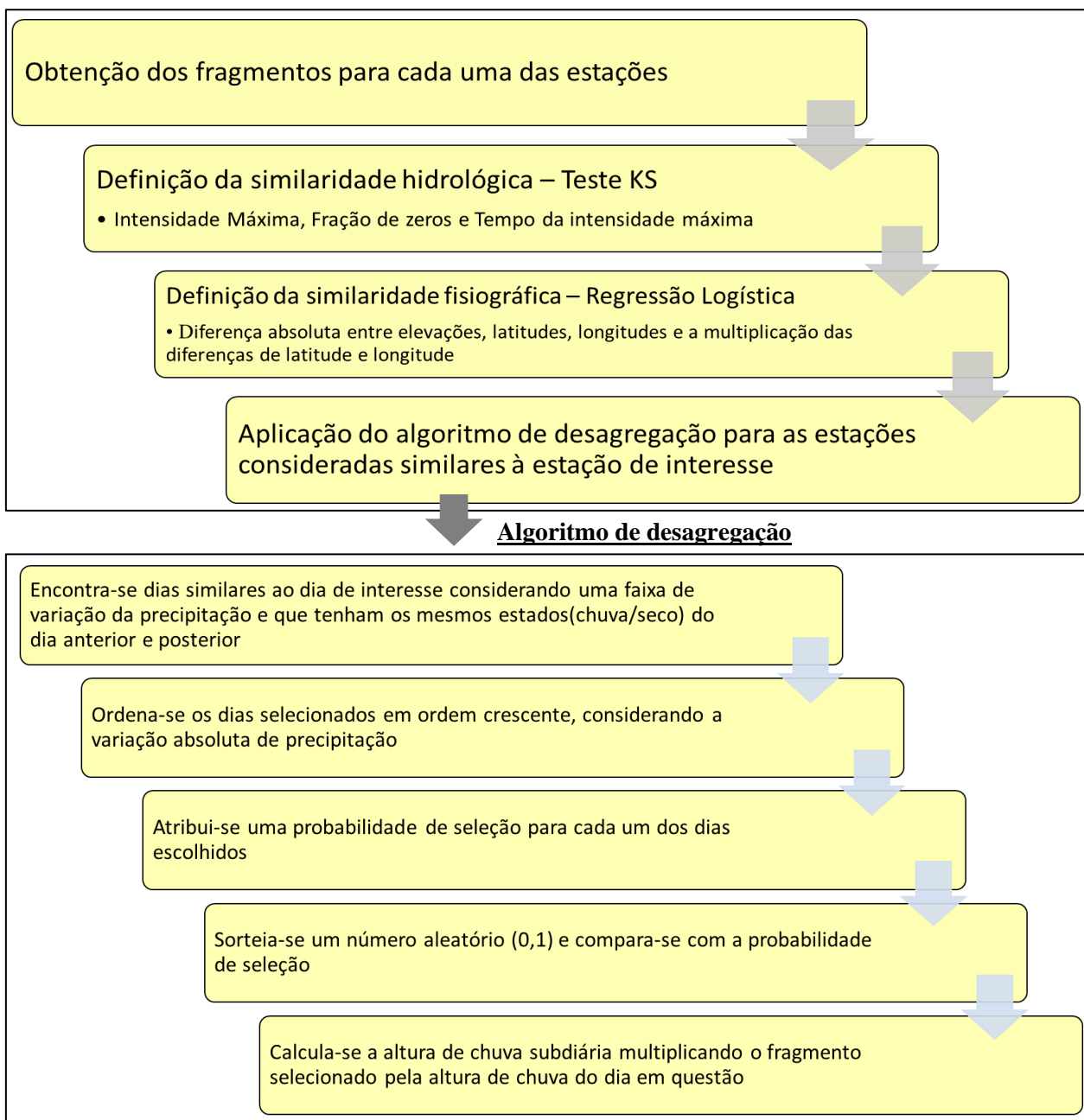


realizar a abordagem regional, ou seja, foram utilizadas todas as estações para amostragem dos fragmentos.

Após a definição do número de vizinhos que resulta em um melhor desempenho do modelo, avaliou-se o limite da variação absoluta e a premissa de um número máximo de vizinhos ( $k$ ). Em relação a esses parâmetros, os autores desenvolvedores do algoritmo não avaliaram valores diferentes, apenas indicaram a utilização de um limite de 10% e um número máximo de 10 vizinhos. Esses valores foram escolhidos pois, de acordo com Westra *et al.* (2012), garantem uma tolerância para a precipitação diária e uma quantidade significativa de variabilidade amostral ao modelo. Visando avaliar o impacto desses parâmetros, foram analisados diferentes limites a saber: 5% e 10%, além do impacto da utilização de um número máximo de vizinhos ( $k = 10$ ) no modelo.

Em relação à janela móvel, Westra *et al.* (2012) consideraram o seu tamanho igual a 15 dias, pois consideram que esse é um valor adequado para garantir que os fragmentos fossem amostrados sempre do mesmo período do ano, preservando assim a sazonalidade. Considerando que no gerador diário foi observado por Costa (2015) que uma janela pequena possui dificuldades de reproduzir as estatísticas da série observada e que uma janela grande apresenta maior esforço computacional sem agregar melhoras significativas, definiu-se por manter a janela de 15 dias proposta por Westra *et al.* (2012).

A Figura 4.6 resume o procedimento metodológico do gerador subdiário. Primeiramente são calculados os fragmentos, em seguida a definição das estações que serão utilizadas no algoritmo de desagregação. Por fim, o algoritmo é aplicado para a desagregação das séries diárias simuladas por meio do gerador diário apresentado no item anterior.



**Figura 4.6** – Resumo da metodologia do gerador subdiário

Frente ao exposto, foi realizada inicialmente uma análise de sensibilidade para avaliação dos parâmetros. Após a definição daqueles que proporcionam um melhor desempenho do modelo, foi realizada a etapa de validação.

#### **4.4 Etapa 4 – Validação e avaliação da eficiência do modelo**

Após o desenvolvimento do modelo, esse foi validado na estação de Caeté, visto que essa estação possui uma série subdiária relativamente longa (23 anos) quando comparada com as

demais estações (média 6 anos). Devido à dificuldade de se encontrar outras séries longas para validação, apenas essa estação foi utilizada. O desempenho do gerador foi avaliado para diferentes escalas subdiárias, tanto para a simulação de eventos regulares quanto extremos, comparando as estatísticas das séries simuladas com as séries observadas dos pluviógrafos, como: estatísticas subdiárias para uma determinada duração, como média, desvio padrão, coeficiente de assimetria e alturas máximas de precipitação.

Outros aspectos também serão avaliados, como: a identificação das principais características para preservação da distribuição conjunta, ou seja, se alguma das características hidrológicas ou fisiográficas tem um maior impacto que as demais no modelo; avaliação da capacidade de reprodução de extremos frente a outras abordagens; e avaliação da relação entre os extremos de precipitação diária e aqueles em escala subdiária.

## 5 RESULTADOS E DISCUSSÃO

### 5.1 Avaliação dos dados hidrológicos

A Tabela 5.1 apresenta informações sobre as estações utilizadas neste trabalho, tais como latitude, longitude e elevação, inclusive da estação de interesse, estação de Caeté (código 01943010), em negrito. As elevações das estações possuem média de El. 700 m, variando entre El. 1326 m e El. 547 m. Em relação às diferenças em latitude e longitude da estação de interesse e as demais estações, observa-se que as estações possuem uma variação em latitude e longitude entre 0.05 e 2.92 graus com a estação de interesse, o que corresponde a aproximadamente uma variação de 5 a 325 km. Em relação a elevação, as diferenças encontram-se na faixa de 5 a 713 m.

**Tabela 5.1** – Informações fisiográficas dos postos pluviográficos e estação de interesse

<b>COD.</b>	<b>NOME</b>	<b>Sub-bacia</b>	<b>Latitude</b>	<b>Longitude</b>	<b>Elevação</b>
<b>01943010</b>	<b>CAETE (Estação de interesse)</b>	<b>41</b>	<b>19° 54' 10"</b>	<b>43° 39' 59"</b>	<b>825</b>
02043002	LAGOA GRANDE (MMV)	41	20° 10' 45"	43° 56' 34"	1326
01943022	CAIXA DE AREIA	41	19° 57' 02"	43° 54' 10"	1159
01943009	VESPASIANO	41	19° 41' 13"	43° 55' 14"	677
01944009	PEDRO LEOPOLDO	41	19° 38' 04"	44° 03' 09"	730
01946009	SÃO GOTARDO	41	19° 18' 51"	46° 02' 39"	1086
01943035	VAU DA LAGOA	41	19° 13' 05"	43° 35' 17"	1085
01844010	PONTE DO LICÍNIO-JUSANTE	41	18° 40' 20"	44° 11' 36"	547
01844009	PRESIDENTE JUSCELINO-JUSANTE	41	18° 38' 41"	44° 03' 02"	576
01843002	GOUVEA	41	18° 27' 56"	43° 44' 35"	1106
01844001	SANTO HIPOLITO (ANEEL/CEMIG)	41	18° 18' 21"	44° 13' 32"	530
01845021	CANOEIROS	41	18° 02' 17"	45° 31' 23"	796
02044007	ENTRE RIOS DE MINAS	40	20° 39' 37"	44° 04' 18"	871
02044042	CARMO DA MATA (ETA-COPASA)	40	20° 33' 45"	44° 52' 02"	854
02043013	CONGONHAS-LINIGRAFO	40	20° 31' 06"	43° 50' 08"	871
02045012	PIUM-I	40	20° 27' 33"	45° 56' 38"	809
02045002	IGUATAMA	40	20° 10' 12"	45° 42' 56"	639
02045013	SANTO ANTONIO DO MONTE	40	20° 05' 03"	45° 17' 48"	969
02044052	JARDIM	40	20° 02' 50"	44° 24' 32"	779

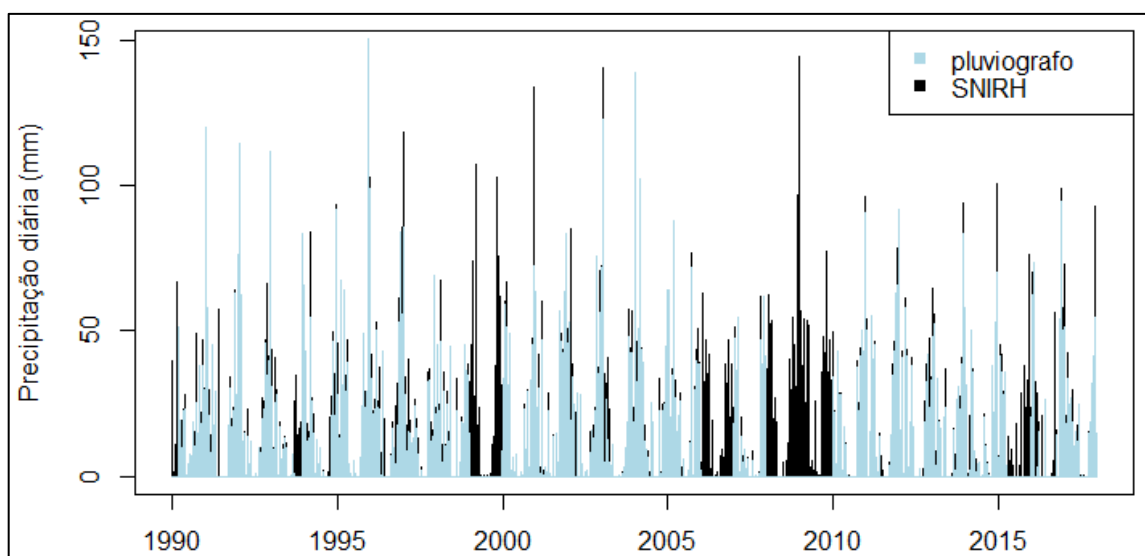
(continua)

**Tabela 5.1** – Informações fisiográficas dos postos pluviográficos e estação de interesse (continuação)

<b>COD.</b>	<b>NOME</b>	<b>Sub-bacia</b>	<b>Latitude</b>	<b>Longitude</b>	<b>Elevação</b>
01944027	JUATUBA	40	19° 57' 20"	44° 20' 03"	724
01944004	PONTE NOVA DO PARAOPEBA	40	19° 56' 57"	44° 18' 19"	708
01944021	VELHO DA TAIPA	40	19° 41' 31"	44° 55' 56"	636
01944049	PAPAGAIOS	40	19° 25' 38"	44° 43' 11"	741
01845004	LAGOA DO GOUVEIA	40	18° 49' 59"	45° 50' 26"	1038
01944062	FAZENDA SANTA RITA	40	19° 58' 58"	44° 29' 32"	820
02044021	ALTO DA BOA VISTA	40	20° 06' 20"	44° 24' 04"	905
02044024	FAZENDA CURRALINHO	40	20° 00' 27"	44° 19' 52"	786
02044041	FAZENDA LARANJEIRAS	40	20° 06' 08"	44° 29' 05"	895
02044054	SERRA AZUL	40	20° 05' 12"	44° 25' 38"	817
01841011	TUMIRITINGA	56	18° 58' 15"	41° 38' 30"	137
01940009	PANCAS	56	19° 13' 51"	40° 50' 07"	112
01940020	CALDEIRAO	56	19° 57' 17"	40° 44' 30"	694
01941005	BARRA DO CUIETE- JUSANTE	56	19° 03' 42"	41° 31' 59"	143
01941006	ASSARAI-MONTANTE	56	19° 35' 39"	41° 27' 29"	156
01941012	ALTO RIO NOVO	56	19° 03' 29"	41° 01' 39"	535
01942008	DOM CAVATI	56	19° 22' 26"	42° 06' 07"	319
01942030	CENIBRA	56	19° 19' 40"	42° 23' 51"	225
01942031	CACHOEIRA DOS OCULOS-MONTANTE	56	19° 46' 36"	42° 28' 35"	248
01942032	NAQUE VELHO	56	19° 11' 17"	42° 25' 20"	205
01943002	CONCEICAO DO MATO DENTRO	56	19° 00' 51"	43° 26' 48"	624
02043010	PIRANGA	56	20° 41' 17"	43° 18' 02"	608

Para avaliar preliminarmente a qualidade dos dados em escala diária, agregou-se os valores de precipitação subdiária em precipitação diária para realizar a comparação com a série diária obtida do SNIRH. A série para a estação de Caeté no SNIRH apresenta dados de precipitação diária observados entre 1942 e 2018, totalizando 77 anos de dados. De modo a realizar a comparação mencionada, foi elaborado um gráfico de séries temporais, selecionando apenas os dados da série diária do SNIRH para o mesmo período, conforme apresentado na Figura 5.1. Observa-se pelos grandes intervalos sem precipitação presentes no gráfico que a série provinda do pluviógrafo possui alguns anos com uma grande quantidade de dados faltantes. Por esse motivo, 5 anos da série do pluviógrafo foram removidos, a saber 1999, 2006, 2008, 2009 e 2015, totalizando 23 anos de dados de precipitação na série final. Observa-se também que a

série do pluviógrafo que foi agregada não apresenta grandes discrepâncias quando comparada com a série diária do SNIRH, o que é importante considerando o aspecto de qualidade dos dados.



**Figura 5.1** – Comparação da série temporal de precipitação diária do SNIRH e a série subdiária agregada do pluviógrafo para a estação de Caeté

Continuando a avaliação dos dados, obteve-se a série de máximos anuais diários para os 23 anos, visando agora avaliar os eventos máximos que serão utilizados na estimação da distribuição de extremos. Foram realizados testes de significância para avaliar a aleatoriedade, homogeneidade, independência e estacionariedade dos dados. Dos testes realizados, a hipótese nula de independência foi rejeitada considerando um nível de significância de 5%.

Visto o resultado de rejeição de um dos testes de significância e que uma série de 23 anos é relativamente curta para estimação de máximos de precipitação, optou-se por utilizar os máximos anuais da série de precipitação diária obtida do SNIRH (77 anos), para estimação dos parâmetros da distribuição de extremos. Foram realizados novamente os testes de significância para os máximos anuais desse grupo de dados e obteve-se, como resultado a não rejeição da hipótese nula em todos testes mencionados anteriormente, considerando um nível de significância de 5%. A utilização de uma série mais longa de máximos anuais (77 anos) oriundas dos dados do SNIRH proporciona, ao menos em teoria, uma amostra mais representativa para estimação dos parâmetros da distribuição LN4. Para a geração das séries diárias e desagregação das mesmas utilizou-se a série de 23 anos do pluviógrafo, de modo a promover a etapa de validação e avaliação dos modelos.

## 5.2 Gerador estocástico diário

### 5.2.1 Distribuição *a priori* para o limite superior

Não foi encontrada uma referência de PMP de 1 dia utilizando dados hidrometeorológicos para a estação de interesse, a estação de Caeté, sendo assim, utilizou-se a estimativa de uma estação mais próxima àquela de interesse. O valor encontrado para tal quantidade é de 397 mm em uma estação em Mariana, pertencente à Sub-bacia 56 do rio Doce (PINHEIRO, 2011). De acordo com o procedimento especificado no item 4.2.1, a distribuição gama será empregada para modelar as incertezas com relação ao limite superior. Dessa forma, torna-se necessário estimar o coeficiente de variação regional das PMPs e a probabilidade de não-excedência da PMP local.

Conforme determinado por Costa (2015), o coeficiente de variação estimado para o conjunto de 118 estimativas de PMP estatística é igual a 0,154. Considerando a estimativa meteorológica e a posição de plotagem de Weibull, foi estimada a frequência de não-excedência empírica para a estimativa local de PMP, sendo essa igual a 0,186. Portanto, a distribuição *a priori* eliciada para o limite superior é a  $\sim \text{GAMA}(42,166; 0,091)$ , com as principais características apresentadas na Tabela 5.2. Formalmente a plotagem de Weibull é definida como:

$$q_i = \frac{i}{N + i} \quad (5.1)$$

em que  $i$  é a posição na amostra ordenada e  $N$  é o tamanho da amostra.

**Tabela 5.2** – Parâmetros e características das distribuições *a priori* do limite superior

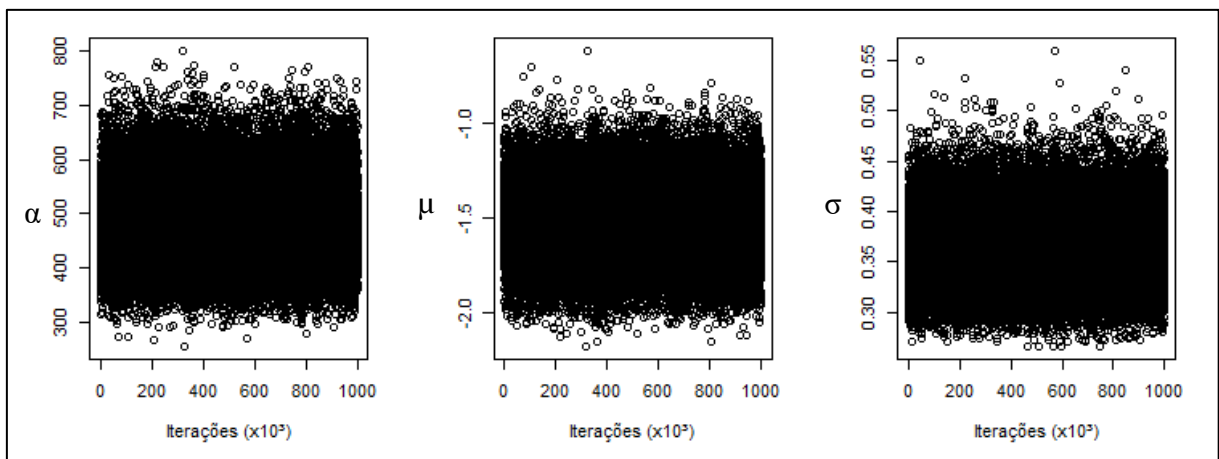
Distribuição <i>a priori</i>	$\rho_\alpha$	$\beta_\alpha$	Média	Mediana	$CV_\alpha$	DP
Gama	42,166	0,091	463,4	459,7	0,154	71,4

### 5.2.2 Distribuição *a posteriori* para os parâmetros da distribuição LN4

Como apresentado pela Equação (4.3), a distribuição *a posteriori* dos parâmetros é proporcional ao produto da função de verossimilhança pela distribuição conjunta *a priori*, onde  $p(x|\Theta)$  é a função de verossimilhança,  $p(\alpha)$ ,  $p(\mu)$  e  $p(\sigma)$  são, respectivamente, as distribuições *a priori* para o limite superior, para o parâmetro de posição e para o parâmetro de escala. Para os

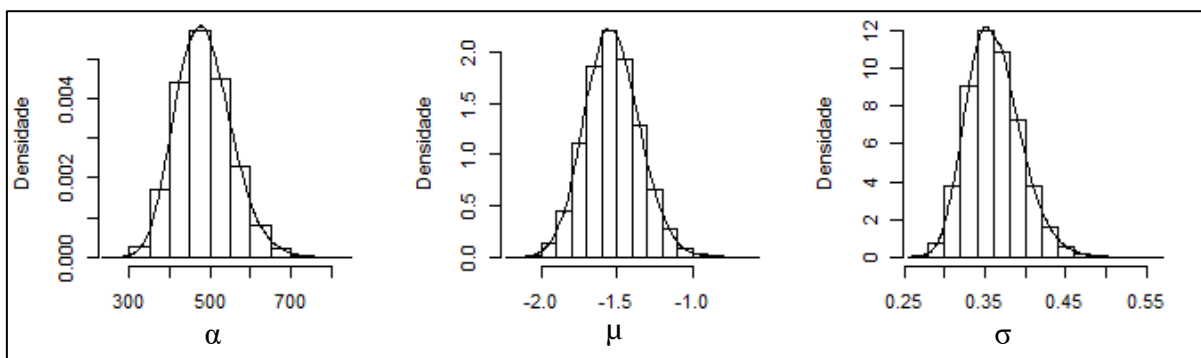
parâmetros  $\mu$  e  $\sigma$ , a ausência de relação direta dos mesmos com o fenômeno físico fez com que fossem empregadas distribuições *a priori* não informativas. Uma vez que  $\mu$  pode assumir qualquer valor real e  $\sigma$  é sempre positivo, admitiu-se uma distribuição normal não informativa para o primeiro e uma distribuição gama não informativa para o segundo, ou seja,  $\mu \sim \text{NORMAL}(1,0; 1,0 \times 10^{-6})$  e  $\sigma \sim \text{GAMA}(1,0; 1,0 \times 10^{-8})$ .

Desse modo, para estimação dos parâmetros, considerando a série de máximos anuais de 77 anos da estação de Caeté, empregou-se o software OpenBUGS (LUNN *et al.*, 2009) para realização das simulações numéricas. Conforme preconizado por Costa (2015), foram considerados neste trabalho um *burn-in* de 50.000 e um *lag* de 20 para obtenção de uma amostra final com auto correlação igual a zero, com tamanho de 50.000 para cada um dos parâmetros, sendo tal valor considerado suficiente para caracterizar a variabilidade de suas distribuições marginais *a posteriori*. A Figura 5.2 mostra a variação dos valores dos parâmetros do modelo LN4 ao longo da simulação, após o descarte dos valores de *burn-in* e aplicação do *lag*. Nessa figura é possível observar a homogeneidade dos valores, sem a presença de tendências ou alterações na variância, dentro da região do domínio de cada parâmetro. Já a Figura 5.3 apresenta os histogramas das densidades de cada um dos parâmetros, com a linha contínua representando a distribuição *a posteriori* ajustada ao conjunto de estimativas.



**Figura 5.2** – Variação dos valores dos parâmetros do modelo LN4 ao longo da simulação





**Figura 5.3** – Densidade e histograma dos valores dos parâmetros do modelo LN4 ao longo da simulação

A Tabela 5.3 mostra os resultados *a posteriori* para os parâmetros  $\alpha$ ,  $\mu$  e  $\sigma$ . Visto que não foi possível estabelecer uma distribuição *a priori* informativa para os parâmetros  $\mu$  e  $\sigma$ , e considerando a falta de relações físicas entre os parâmetros mencionados e as características hidrometeorológicas da bacia, analisar os resultados *a posteriori* constitui uma tarefa complexa. Pode-se observar pela Figura 5.3 que a distribuição *a posteriori* desses dois parâmetros é unimodal, sendo aproximadamente simétrica para o parâmetro  $\mu$  e levemente assimétrica à direita para parâmetro  $\sigma$ , com assimetria positiva. Adicionalmente, as medianas estão bem próximas da média, como pode ser visto pela Tabela 5.3, confirmando a tendência de simetria.

Já para o parâmetro  $\alpha$ , observa-se que a distribuição *a posteriori* também é unimodal aproximadamente simétrica, com a estimativa de  $\alpha$  sendo significativamente superior à estimativa pontual de PMP, o que mostra que a utilização desse valor de modo determinístico é inadequada. Além disso, para valores consideravelmente superiores ao limite superior - acima de 800 mm, por exemplo - o modelo atribui probabilidades extremamente reduzidas. Desse modo, a probabilidade de estimação de valores fisicamente implausíveis é pequena.

**Tabela 5.3** – Características das distribuições *a posteriori* dos parâmetros do modelo LN4

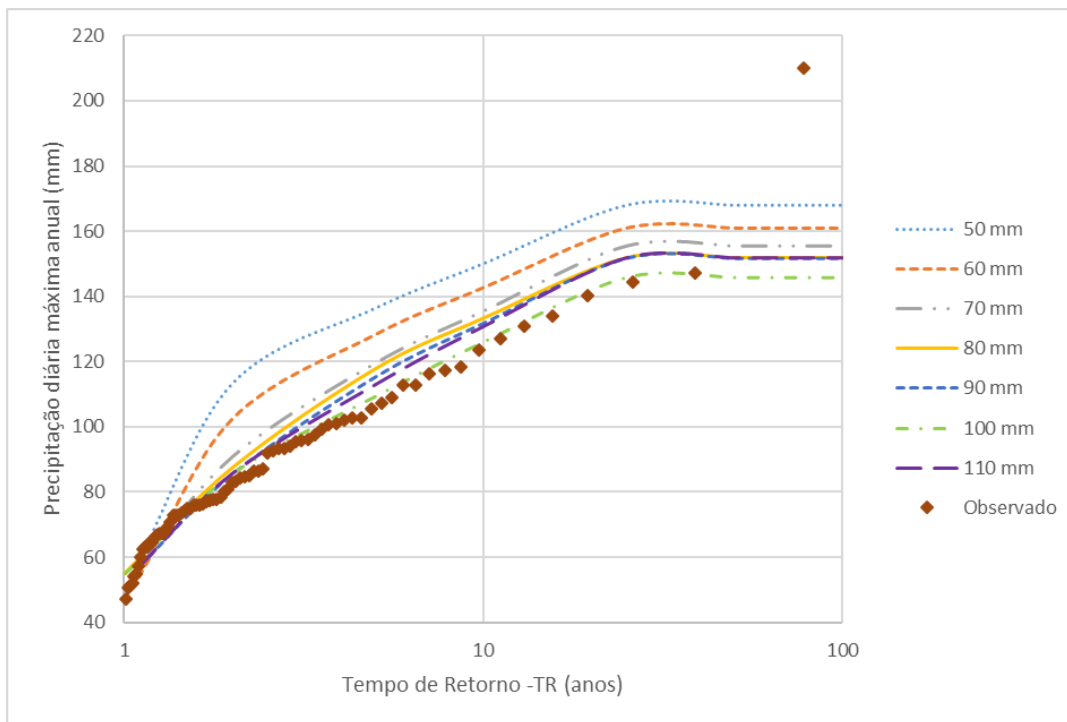
Parâmetro	Média	Mediana	DP	CV	95% HPD
$\alpha$	487,0	482,1	68,89	0,14	(366,5;632,2)
$\mu$	-1,533	-1,537	0,179	0,117	(-1,874;-1,174)
$\sigma$	0,361	0,359	0,033	0,091	(0,304;0,432)

95% HPD – Intervalo de credibilidade

### 5.2.3 Definição do limiar entre os eventos regulares e extremos

De posse das estimativas pontuais dos parâmetros da distribuição LN4 e empregando o tamanho da janela de reamostragem do módulo não paramétrico igual a 14 dias, definiu-se o limiar entre chuvas convencionais e extremas. Considerando os aspectos discutidos nos itens 3.1.5 e 4.2, para essa definição, utilizou-se de funções-objetivo, como a curva de quantis de precipitações diárias máximas anuais, precipitações médias mensais e anuais.

Foram testados limites variando entre 50 mm e 110 mm, gerando 1.000 séries de 23 anos (tamanho da amostra subdiária na estação de interesse), calculando os valores médios das estatísticas. A Figura 5.4 apresenta as curvas de quantis de precipitações diárias máximas anuais. É possível observar que para o valor de 100 mm ocorre um razoável ajuste ao conjunto de registros observados. O ajuste não ocorre para os demais valores, devido ao uso excessivo da distribuição LN4 ao longo das simulações, o que proporciona a geração de mais eventos extremos do que verificado nos dados observados. É importante ressaltar que o máximo anual observado de aproximadamente 210 mm aparenta, à primeira vista, ser um *outlier* do conjunto de dados. Entretanto, não há nenhum registro que confirme que esse valor venha a ser um erro de medição. Provavelmente, pelo tamanho reduzido da amostra, há uma associação de uma probabilidade empírica incorreta referente a essa observação, a qual deve estar associada a um tempo de retorno mais elevado do que aquele associado à sua posição na amostra. Essa observação também é discutida no Capítulo 8, Item 8.1.2 do livro de Naghettini & Pinto (2007), sem conclusões a respeito dessa questão.



**Figura 5.4** – Curvas de quantis de precipitações diárias máximas anuais para diferentes limiares entre eventos regulares e extremos

Visando avaliar influências na variância para durações maiores, elaborou-se a Tabela 5.4, que apresenta a precipitação média anual para cada limiar e a variação em porcentagem em relação ao observado. Observa-se que o limiar de 50 mm é o que mais se diferencia dos demais, sendo que, a partir do limiar de 70 mm, a diferença fica menor que 1%.

**Tabela 5.4** – Precipitação média anual e variação em relação ao observado para cada limiar

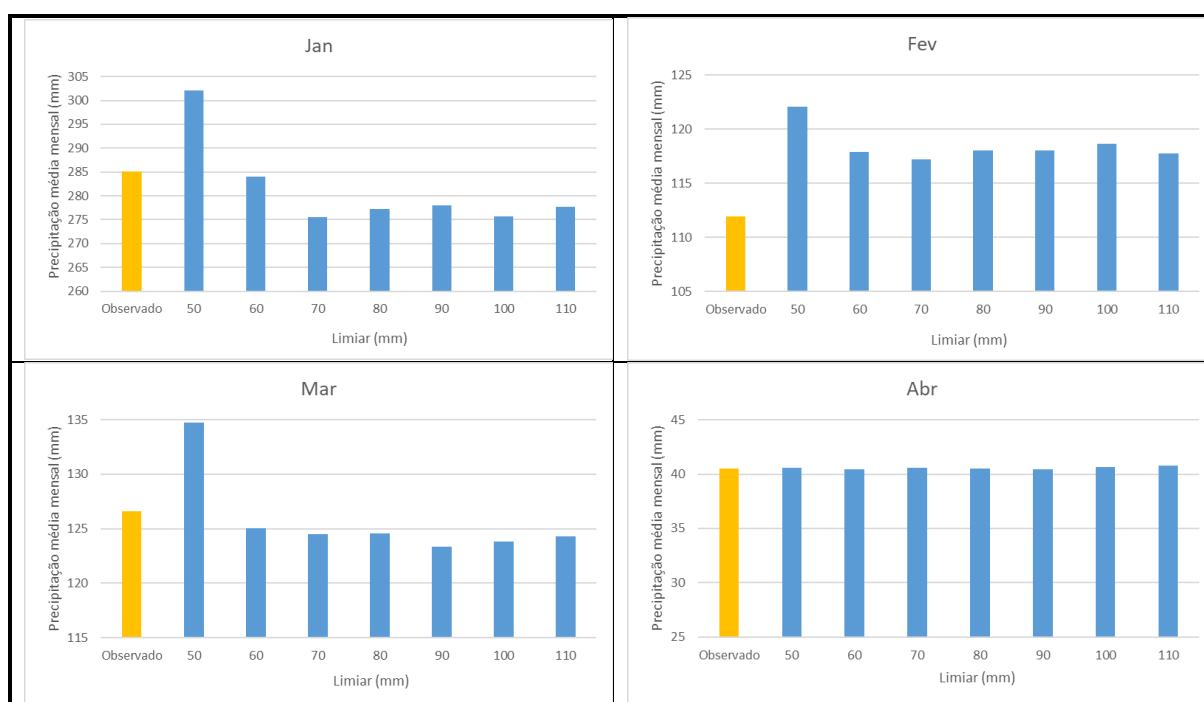
	Anual (mm)	Varição anual
<b>Observado</b>	1262,395	-
<b>Limiar de 50 mm</b>	1350,874	7,01%
<b>Limiar de 60 mm</b>	1292,94	2,42%
<b>Limiar de 70 mm</b>	1268,971	0,52%
<b>Limiar de 80 mm</b>	1268,945	0,52%
<b>Limiar de 90 mm</b>	1264,776	0,19%
<b>Limiar de 100 mm</b>	1261,831	-0,04%
<b>Limiar de 110 mm</b>	1262,767	0,03%

Ainda nesse contexto, construiu-se a Fig. 5.5, que apresenta as precipitações médias mensais para os diferentes limiares. Observa-se que apenas o limiar de 50 mm não reflete apropriadamente

as precipitações observadas, assim como mostrado pela Tabela 5.4. Para os demais limiares, nos meses entre abril e outubro, ou seja, meses mais secos do ano, o comportamento é bem similar. Já nos meses mais chuvosos, novembro a fevereiro, o comportamento varia bastante dependendo do limiar.

O limiar de 60 mm funciona bem para os meses chuvosos, com exceção de dezembro. Por ser um limiar mais baixo, há uma geração excessiva de eventos considerados extremos, superestimando assim a precipitação média mensal. Avaliando o limiar de 100 mm, que se adaptou melhor ao conjunto de máximos anuais observados, observa-se que o mesmo apresenta valores inferiores ao de 60 mm para os meses chuvosos, principalmente para o mês dezembro. A utilização de um limiar superior diminui, assim, a geração de eventos extremos excessivos, apresentando melhor desempenho em meses mais chuvosos.

Uma desvantagem da utilização de um limiar tão alto é que esse impede a geração de chuvas extremas entre os meses de abril a outubro. Entretanto, em virtude das características climáticas e de sazonalidade da região de estudo, a probabilidade de ocorrência de um máximo anual fora do intervalo entre novembro e março é muito baixa. Para preservar o comportamento da cauda superior das precipitações diárias e diminuir impactos nas médias mensal e anual para durações maiores, definiu-se pela escolha do limiar de 100 mm para o gerador estocástico diário.



**Figura 5.5** – Precipitações médias mensais para diferentes limiares entre eventos regulares e extremos (continua)

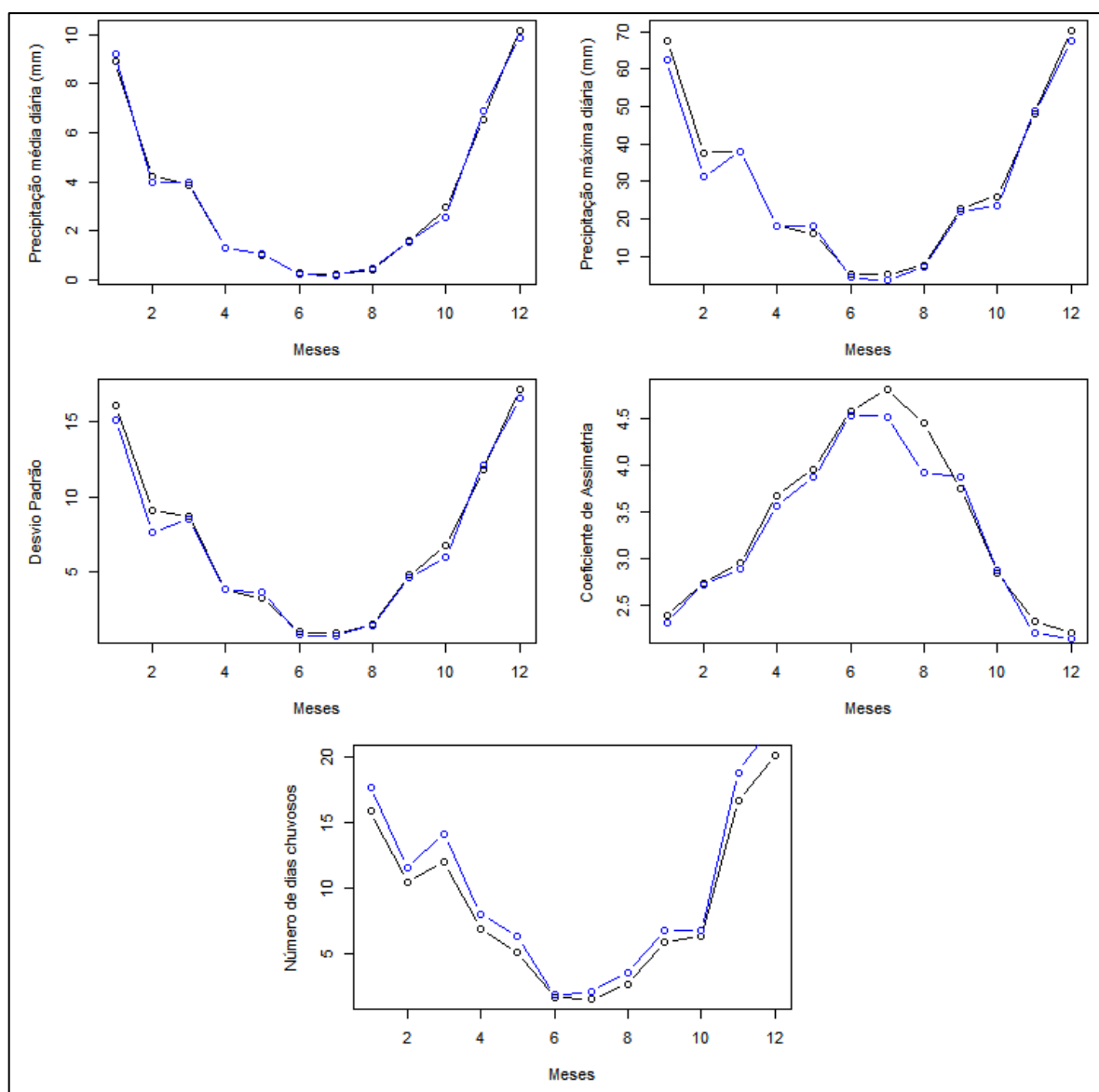


**Figura 5.5** - Precipitações médias mensais para diferentes limiares entre eventos regulares e extremos

De posse das estimativas pontuais dos parâmetros da distribuição LN4 e após a definição da janela de amostragem (14 dias) e do limiar entre os eventos regulares e extremos (100 mm), o algoritmo do gerador estocástico de precipitação diária foi utilizado para a geração das 1.000 séries de precipitação diária para posterior emprego no gerador subsidiário.

### 5.2.4 Avaliação do desempenho do gerador estocástico diário

Para avaliar o desempenho do gerador, utilizou-se as estatísticas diárias média, máxima, desvio padrão, coeficiente de assimetria e números de dias chuvosos para comparação entre as séries simuladas e as séries observadas. A Figura 5.6 apresenta essa comparação, com os valores observados em azul e os simulados em preto. É possível observar que o modelo utilizado reproduz de maneira adequada a média, os máximos e o desvio padrão da chuva diária, para todos os meses do ano, confirmando a escolha adequada do limiar e do tamanho da janela de reamostragem. Em relação à assimetria, o modelo apresenta uma maior variação para os meses de julho e agosto.

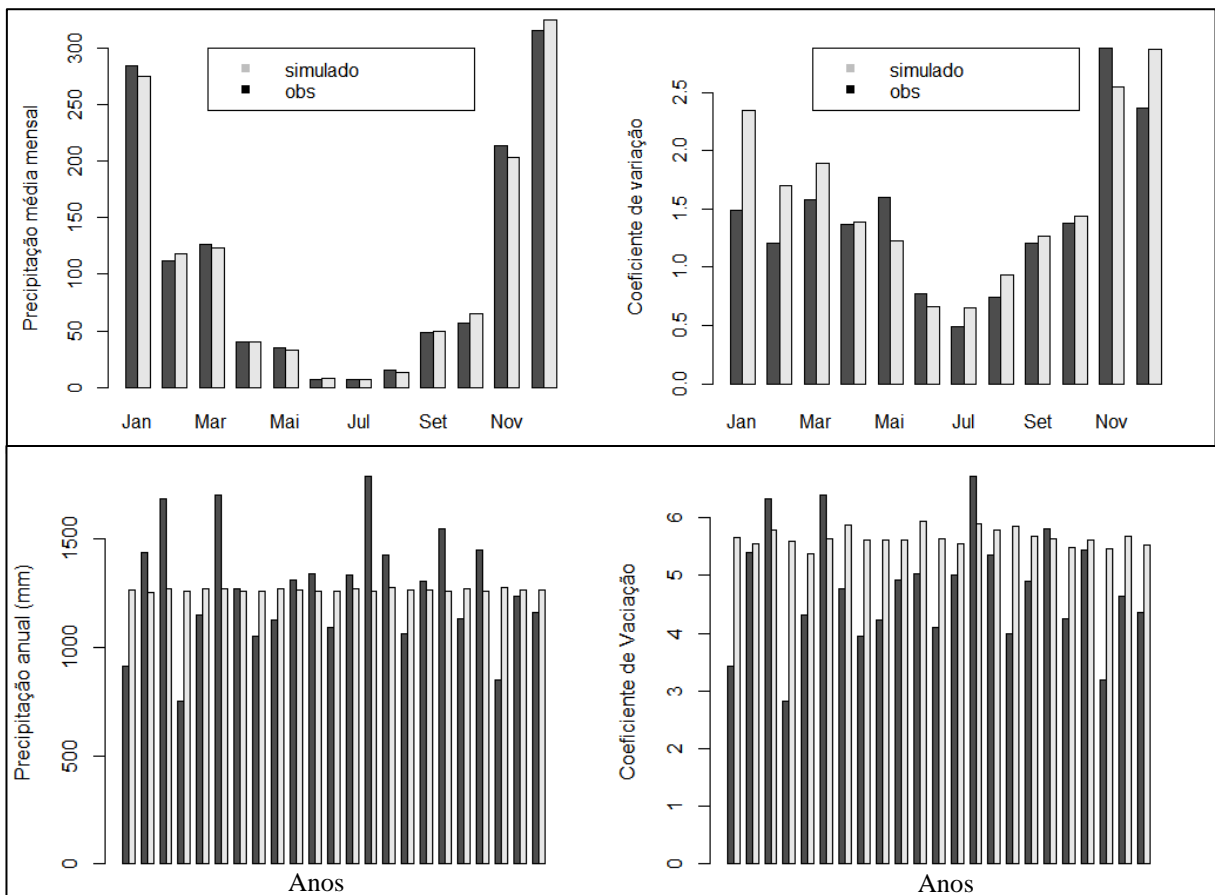


**Figura 5.6** – Comparação das estatísticas entre as séries simuladas (preto) e as séries observadas (azul)

Com relação ao número médio de dias chuvosos, pode-se observar um viés de subestimação, mais evidente nos meses da estação chuvosa (novembro-março). Já Costa (2015) obteve o viés mais evidente na estação seca. Conforme apresentado por esse autor, o viés pode estar relacionado à construção das matrizes de probabilidade de transição em uma base diária, em lugar da escala mensal usualmente empregada.

Visando analisar o desempenho do modelo em escalas temporais mais longas, como mensais e anuais, foi elaborada a Figura 5.7. Essa análise é importante já que modelos paramétricos comumente reproduzem as estatísticas mencionadas anteriormente para a escala diária, porém apresentam tendência de subestimação da variância para escalas mais longas (COSTA, 2015).

Como pode-se observar pela Figura 5.7 o gerador foi capaz de reproduzir de forma apropriada as médias e variância mensais, embora apresentando uma tendência de superestimação. Já para a escala anual, observa-se uma tendência de estabilização da precipitação anual do modelo em torno de 1250 mm. Isso se deve ao fato de que estão sendo analisadas as médias das 1.000 séries para cada ano. Entretanto, quando se analisa o coeficiente de variação, observa-se que há uma variação das precipitações anuais a cada simulação, apresentando o mesmo comportamento de superestimação indicado pelas estatísticas mensais.



**Figura 5.7** – Comparação das médias e coeficiente de variação mensal e anual entre as séries simuladas (cinza) e as séries observadas (preto)

### 5.3 Gerador estocástico subdiário

#### 5.3.1 Definição da similaridade hidrológica e fisiográfica

Após a obtenção e conversão dos dados das 40 postos pluviográficos para as durações escolhidas (60, 180, 360 e 720 minutos), foram calculados os fragmentos e os atributos hidrológicos (intensidade máxima, tempo da intensidade máxima e fração de zeros) para cada uma das estações, conforme indicado na Seção 4.3 deste trabalho.

Por meio da aplicação do teste KS de duas dimensões e duas amostras, foram definidas as estações que são estatisticamente similares, considerando a distribuição conjunta entre cada atributo hidrológico e os volumes de precipitação diária. Como premissa, considerou-se um nível de significância de 5% e separou-se o modelo em estações do ano para levar em conta as influências sazonais.



Aplicando o teste para as estações escolhidas, observou-se que esse funcionou apropriadamente para o atributo de intensidade máxima. Em relação ao atributo tempo da intensidade máxima, o comportamento apresentado indicou uma tendência de diminuição da similaridade com o aumento da duração. Isso pode ser explicado pelo comportamento dos eventos de precipitação, que em durações menores podem estar relacionados a processos convectivos, apresentando uma similaridade de ocorrer no mesmo horário do dia, e em maiores durações, tendem a estar relacionados a processos frontais, que envolvem massas de ar e sistemas mais complexos, não correlacionados diretamente ao horário do dia.

Já para o atributo fração de zeros, o comportamento apresentado indicou similaridade apenas para a estação do ano inverno (junho, julho e agosto). Isso pode ser explicado já que no inverno há poucos dias chuvosos, facilitando a identificação de correlações. Ressalta-se que Westra *et al.* (2012) também encontraram problemas com o atributo fração de zeros, apresentando baixas probabilidades de similaridade. Entretanto, os autores só apresentam os resultados para as durações de 1 hora e 6 minutos, dificultando a comparação para durações maiores.

O próximo passo, conforme o método proposto, é a definição da similaridade considerando as características fisiográficas, quais sejam, a diferença absoluta entre elevações, latitudes, longitudes e a multiplicação das diferenças de latitude e longitude. As informações fisiográficas das estações foram apresentadas no item 4.1.

Foram calculadas as diferenças absolutas das características fisiográficas entre os pares das estações, com exceção da estação de interesse. As características fisiográficas entre a estação de interesse e demais estações serão utilizadas após a obtenção dos coeficientes de regressão. Assim, utilizando a resposta binária obtida pelo teste KS, aplicou-se a Regressão Logística para cada uma das características hidrológicas. A Tabela 5.5 apresenta os coeficientes do modelo de regressão encontrados para cada característica hidrológica para a duração de 60 minutos, considerando a sazonalidade, onde DJF (dezembro, janeiro e fevereiro) refere-se à estação do ano verão, MAM (março, abril e maio) refere-se à estação do ano outono, JJA (junho, julho e agosto) refere-se à estação do ano inverno e SON (setembro, outubro e novembro) refere-se à estação do ano primavera. Para as demais durações, ver Tabelas I.1, I.2 e I.3 do Apêndice I.

**Tabela 5.5** – Coeficientes do modelo de regressão para duração de 60 minutos

<b>Máxima intensidade</b>					
<b>Estação do ano</b>	<b>Intercepto (<math>\beta_0</math>)</b>	<b>Latitude (<math>\beta_1</math>)</b>	<b>Longitude (<math>\beta_2</math>)</b>	<b>LatxLong (<math>\beta_3</math>)</b>	<b>Elevação (<math>\beta_4</math>)</b>
<b>DJF</b>	-0,3816	-0,007	-0,2219	0,0426	-0,0004
<b>MAM</b>	-1,0302	-0,2184	-0,2250	0,0841	0,0001
<b>JJA</b>	0,4919	-0,1374	-0,3586	-0,0528	-0,0007
<b>SON</b>	-0,0017	0,0060	-0,1897	0,1301	-0,0009

<b>Tempo da intensidade máxima</b>					
<b>Estação do ano</b>	<b>Intercepto (<math>\beta_0</math>)</b>	<b>Latitude (<math>\beta_1</math>)</b>	<b>Longitude (<math>\beta_2</math>)</b>	<b>LatxLong (<math>\beta_3</math>)</b>	<b>Elevação (<math>\beta_4</math>)</b>
<b>DJF</b>	-4,7091	-0,4951	-0,2871	0,0463	-0,0005
<b>MAM</b>	-3,2334	0,1207	-0,6331	-0,1007	0,0001
<b>JJA</b>	-0,5423	-0,1544	-0,3661	-0,1239	-0,0008
<b>SON</b>	-1,9019	-0,3080	-0,0247	-0,1183	-0,0011

<b>Fração de zeros</b>					
<b>Estação do ano</b>	<b>Intercepto (<math>\beta_0</math>)</b>	<b>Latitude (<math>\beta_1</math>)</b>	<b>Longitude (<math>\beta_2</math>)</b>	<b>LatxLong (<math>\beta_3</math>)</b>	<b>Elevação (<math>\beta_4</math>)</b>
<b>DJF</b>	-26,5661	0,0000	0,0000	0,0000	0,0000
<b>MAM</b>	-26,5661	0,0000	0,0000	0,0000	0,0000
<b>JJA</b>	-2,0653	-0,0935	-0,0710	0,0191	-0,0012
<b>SON</b>	-26,5661	0,0000	0,0000	0,0000	0,0000

Devido às baixas correlações entre as estações, a resposta binária para o atributo hidrológico fração de zeros foi de 0 ou seja, não similar, para todas as estações pluviográficas, fazendo com que o modelo de regressão logística não convergisse, conforme observado pela Tabela 5.5, exceto na estação do ano inverno (JJA), pois encontrou-se similaridade. Observa-se que acontece o mesmo problema de convergência para o atributo tempo da intensidade máxima para as durações de 6 hrs e 12 hrs (ver Tabela I.2 e I.3 do Apêndice I). Frente ao exposto, esse comportamento impediu a utilização do atributo fração de zeros no modelo para todas as durações e do atributo tempo da intensidade máxima para duração de 12 horas.

Após a obtenção dos coeficientes de regressão, foram calculadas as características fisiográficas entre a estação de interesse e as demais estações. Considerando ambas informações, as Equações (4.6) e (4.7) foram utilizadas para obtenção das probabilidades das estações serem similares à estação de interesse para cada característica hidrológica, estação do ano e duração. As Tabelas I.4 a I.9 no Apêndice I apresentam os valores de probabilidades de similaridade encontrados para cada característica hidrológica, estação do ano e duração. A Tabela 5.6 apresenta um resumo dos valores encontrados para as 20 estações com maior probabilidade de

similaridade, onde os campos indicados por um traço significam que o modelo de regressão logística não convergiu, apresentando probabilidades iguais a zero.

**Tabela 5.6** – Resumo das probabilidades de similaridade encontradas

<b>Durações</b>	<b>60 min</b>	<b>180 min</b>	<b>360 min</b>	<b>720 min</b>
<b>Intensidade máxima – DJF</b>	[0,45-0,63]	[0,52-0,71]	[0,52-0,71]	[0,52-0,74]
<b>Intensidade máxima – MAM</b>	[0,27-0,34]	[0,43-0,56]	[0,46-0,53]	[0,42-0,51]
<b>Intensidade máxima – JJA</b>	[0,66-0,79]	[0,78-0,87]	[0,76-0,84]	[0,74-0,82]
<b>Intensidade máxima – SON</b>	[0,49-0,65]	[0,64-0,77]	[0,66-0,76]	[0,66-0,78]
<b>Tempo da intensidade máxima - DJF</b>	[0,01-0,04]	-	-	-
<b>Tempo da intensidade máxima - MAM</b>	[0,04-0,06]	[0,01-0,02]	[0,00-0,01]	-
<b>Tempo da intensidade máxima - JJA</b>	[0,36-0,58]	[0,34-0,59]	[0,20-0,40]	-
<b>Tempo da Intensidade Máxima - SON</b>	[0,14-0,19]	[0,01-0,03]	-	-

Observou-se por meio desses resultados que, para o atributo intensidade máxima, para as estações com maior similaridade as probabilidades situam-se entre 0,6 a 0,7, com exceção feita às estações mais chuvosas, nas quais essa probabilidade é menor. Esse comportamento é explicado pelo maior número de dias chuvosos, dificultando a obtenção de correlações entre as estações.

Já para o atributo tempo da intensidade máxima, observa-se que as probabilidades foram diminuindo com a duração, comprovando resultado obtido por meio do teste KS e em conformidade com o mencionado anteriormente acerca da diferença do comportamento dos eventos de precipitação. A estação do ano inverno (JJA) resultou em probabilidades superiores.

Conforme mencionado, como no inverno há poucos dias chuvosos, é facilitada a identificação de correlações, apresentando maior similaridade.

Westra *et al.* (2012) encontraram probabilidades em torno de 0,6 a 0,7 para a duração de 6 minutos, tanto para o atributo hidrológico intensidade máxima, quanto para o atributo tempo da intensidade máxima. Considerando o número menor de estações utilizadas neste trabalho e que para durações maiores é mais difícil de serem obtidas correlações para o atributo tempo da intensidade máxima, os valores encontrados são assumidos apropriados.

Em seguida, considerou-se a média das probabilidades obtidas para cada atributo hidrológico, visando a definição de quais estações serão utilizadas na etapa de desagregação da chuva. A Tabela 5.7 apresenta, para cada uma das estações do ano e duração de 60 minutos, as probabilidades das estações serem similares à estação de interesse. Para as demais durações ver Tabelas I.10, I.11 e I.12 no Apêndice I. Observa-se pela Tabela 5.7 que os valores encontrados para similaridade são maiores para estação de inverno (JJA), conforme os motivos mencionados anteriormente. Outro ponto observado é que, como as probabilidades do atributo tempo da intensidade máxima foram pequenas, exceto para o inverno, ao utilizar a média dos atributos houve uma redução das probabilidades de similaridade. Não obstante, mesmo com probabilidades baixas, utilizou-se esse atributo no modelo, visto que os valores das diferenças entre as estações com maior probabilidade de similaridade ainda são da ordem de magnitude apresentada pelo atributo.

**Tabela 5.7** – Probabilidades de similaridade por estação do ano – Duração de 60 min

DJF		MAM		JJA		SON	
Estação	Prob.	Estação	Prob.	Estação	Prob.	Estação	Prob.
01943009	0,320	01943035	0,199	01843002	0,685	01843002	0,418
02044024	0,315	02043002	0,198	02043013	0,637	01844009	0,382
01943022	0,314	01943022	0,197	01943002	0,624	01844001	0,379
01944009	0,314	01843002	0,191	01844009	0,623	01943002	0,379
01944004	0,311	01943009	0,189	02044007	0,618	02043013	0,378
01944027	0,311	02043013	0,188	01844001	0,611	01943035	0,372
01944062	0,309	01943002	0,188	01943035	0,610	02044007	0,369
02043013	0,306	02043010	0,182	01844010	0,589	01844010	0,365
02044054	0,306	01844009	0,180	02043010	0,588	02043010	0,360
02044052	0,305	01944009	0,179	01944009	0,564	01943009	0,341
02044021	0,296	01844010	0,176	01943009	0,563	01944009	0,340
02044041	0,292	01844001	0,175	02044054	0,532	02044054	0,321
02044007	0,278	02044007	0,175	02044024	0,530	02044024	0,321
01943035	0,276	01944004	0,164	02044052	0,522	01943022	0,320
02043002	0,272	01944027	0,162	02044021	0,520	02043002	0,319
01943002	0,256	02044024	0,160	01944062	0,517	02044021	0,318
02043010	0,254	02044021	0,158	02044041	0,512	02044052	0,317
01944049	0,251	02044052	0,156	01944027	0,512	01944004	0,314
01944021	0,243	02044054	0,154	01944004	0,511	01944027	0,314
02044042	0,235	02044041	0,153	01943022	0,507	02044041	0,314
02045013	0,224	01942032	0,153	02044042	0,498	01944062	0,313
01942031	0,210	01944062	0,150	01944049	0,496	01944049	0,308
01844009	0,207	01942030	0,149	02043002	0,487	02044042	0,308
01843002	0,204	01942031	0,147	01944021	0,431	01944021	0,280
01844010	0,198	01845021	0,144	01845021	0,420	01845021	0,273
02045002	0,185	01944049	0,144	02045013	0,397	01942032	0,267
02045012	0,174	02044042	0,138	01942032	0,361	01942030	0,263
01942030	0,169	01944021	0,134	01942030	0,356	02045013	0,262
01844001	0,166	01942008	0,133	01942031	0,352	01942031	0,258
01942008	0,164	01841011	0,129	02045012	0,347	01942008	0,249
01942032	0,159	01941005	0,123	02045002	0,340	02045002	0,237
01940020	0,145	02045013	0,116	01942008	0,336	02045012	0,235
01946009	0,140	01845004	0,116	01845004	0,317	01845004	0,229
01845021	0,133	01941006	0,108	01946009	0,284	01946009	0,214
01845004	0,132	02045002	0,102	01940020	0,264	01841011	0,212
01941006	0,125	01946009	0,099	01841011	0,247	01941005	0,205
01941012	0,113	01941012	0,098	01941012	0,242	01941006	0,202
01841011	0,104	01940009	0,097	01941005	0,235	01940020	0,199
01941005	0,104	02045012	0,096	01941006	0,232	01941012	0,194
01940009	0,081	01940020	0,069	01940009	0,162	01940009	0,164

### 5.3.2 Calibração do modelo de desagregação

Com as probabilidades de similaridade definidas, os passos do item 4.3.3 foram considerados para obtenção da chuva subdiária da estação de interesse. Após testes preliminares para a duração de 60 min, observou-se que o modelo gastava por volta de 24 horas para desagregação das 1.000 séries de precipitação diária. Assim, devido ao alto custo computacional, a escala do trabalho foi reduzida, conforme mencionado no item 4.3. Foram escolhidas apenas quatro durações neste trabalho, quais sejam, as durações de 60, 180, 360 e 720 minutos, visando identificar diferentes comportamentos ao aumentar a resolução temporal do modelo.

O alto custo computacional impactou também na calibração do modelo, já que é necessária a calibração de vários parâmetros, como o número de estações vizinhas ( $S$ ), o limite da variação absoluta da precipitação e o impacto da utilização do número máximo de vizinhos ( $k$ ). Assim, visto que se torna inviável consumir 24 horas em cada simulação para calibração, diminuiu-se o número de iterações para 100 séries na calibração e o tempo de simulação para 2 horas. Vale ressaltar que Westra *et al.* (2012) também utilizaram em seu trabalho 100 iterações em todas as simulações, devido ao grande número de estações e ao fato de as estações de validação do modelo possuírem mais de 50 anos de dados de precipitação. Li *et al.* (2018) e Pui *et al.* (2012) também utilizaram 100 iterações em seus trabalhos para todas as simulações.

Em relação ao número de estações vizinhas ( $S$ ), analisando os trabalhos de Westra *et al.* (2012), Li *et al.* (2018) e Pui *et al.* (2012), os autores trabalharam com cerca de 200 anos de dados para a amostragem dos fragmentos. Para seguir essa referência neste trabalho, teriam que ser empregadas quase todas as estações no modelo, já que juntas as mesmas somam 240 anos de dados. Após testes preliminares, observou-se o alto custo computacional do modelo ao aumentar o número de vizinhos, nas quais simulações com 5 vizinhos gastam 1,5 horas, acrescentando cerca de 30 minutos adicionais a cada aumento de 5 vizinhos. Assim, para a utilização dos 40 vizinhos, o modelo gasta mais de 12 horas para apenas 100 simulações. Entretanto, também nos testes preliminares, identificou-se o bom comportamento do modelo mesmo utilizando um número baixo de vizinhos e que a utilização de mais de 15 vizinhos não agregava benefícios, ocorrendo, em alguns casos, até em piora dos resultados do modelo. Frente ao exposto, a escolha do número de estações vizinhas ( $S$ ) foi realizada, variando entre 5, 10 e 15 vizinhos, resultando em 30, 60 e 90 anos de dados, respectivamente, considerando que as estações têm em média 6 anos de dados.

Para decisão de qual o conjunto de parâmetros torna o modelo mais eficiente, foram utilizadas as seguintes métricas, considerando apenas os dias chuvosos: (i) média dos máximos anuais; (ii) média; (iii) variância; e (iv) assimetria. Nas Tabelas 5.9 a 5.12 são apresentados os valores observados, os valores simulados e os desvios entre o simulado e o valor observado em porcentagem, para os quais valores negativos indicam subestimação e positivos superestimação do valor observado.

#### Duração de 60 minutos

Os resultados da etapa de calibração para a duração de 60 minutos podem ser vistos na Tabela 5.8. Observa-se que o modelo apresentou valores bem próximos considerando o limite da variação absoluta da precipitação de 5% ou 10%, com o primeiro se comportando um pouco melhor que o segundo. A utilização de um desvio maior implica que mais dias podem ser amostrados para a desagregação, já que só são escolhidos os dias que estão dentro dessa faixa, o que não foi interessante para a duração de 60 minutos.

Em relação ao número de estações vizinhas (S), observa-se que a escolha de 10 vizinhos foi a que se comportou melhor, já que, aparentemente, 5 vizinhos não seriam suficientes e 15 vizinhos prejudicam o desempenho do modelo, amostrando valores de estações que não são tão similares. Observa-se também que o número de estações vizinhas tem um impacto maior no modelo que o parâmetro desvio. Vale ressaltar que os valores encontrados para as métricas ficaram com variação abaixo de 8%, indicando um comportamento apropriado para o modelo.

**Tabela 5.8** – Resultados das simulações de calibração – Parâmetros (S) e Limite – Duração 60 minutos

<b>Duração de 60 min - Limite = 5% - Precipitação em (mm)</b>							
<b>Métricas</b>	<b>Observado</b>	<b>Nº. Vizinhos = 5</b>		<b>Nº. Vizinhos = 10</b>		<b>Nº. Vizinhos = 15</b>	
<b>Média dos máximos anuais</b>	36,98	38,91	5,24%	37,74	2,06%	38,50	4,11%
<b>Média</b>	2,40	2,42	0,98%	2,41	0,64%	2,44	1,61%
<b>Variância</b>	17,58	18,45	4,92%	18,21	3,56%	18,27	3,90%
<b>Assimetria</b>	3,98	4,13	3,85%	4,05	1,62%	4,05	1,82%
<b>Duração de 60 min - Limite = 10% - Precipitação em (mm)</b>							
<b>Métricas</b>	<b>Observado</b>	<b>Nº. Vizinhos = 5</b>		<b>Nº. Vizinhos = 10</b>		<b>Nº. Vizinhos = 15</b>	
<b>Média dos máximos anuais</b>	36,98	38,99	5,45%	38,05	2,91%	38,60	4,39%
<b>Média</b>	2,40	2,46	2,67%	2,46	2,43%	2,46	2,52%
<b>Variância</b>	17,58	18,94	7,71%	18,79	6,88%	18,60	5,78%
<b>Assimetria</b>	3,98	4,08	2,47%	4,02	0,96%	4,02	1,01%

Visto que o modelo se comportou melhor para o limite da variação absoluta da precipitação de 5% e para o número de estações vizinhas igual a 10, utilizou-se desses parâmetros para avaliação do impacto da utilização do número máximo de vizinhos (k). Como pode ser visto pela Tabela 5.9, a utilização do parâmetro k no modelo não agregou uma melhora significativa e, no caso da média, ocasionou uma leve piora.

**Tabela 5.9** – Resultados das simulações de calibração – Parâmetro (k) – Duração 60 minutos

Métricas	Observado	Nº máx. de vizinhos		k = 10	
<b>Média dos máximos anuais</b>	36,98	37,74	2,06%	37,64	1,79%
<b>Média</b>	2,40	2,41	0,64%	2,42	0,94%
<b>Variância</b>	17,58	18,21	3,56%	18,27	3,90%
<b>Assimetria</b>	3,98	4,05	1,62%	4,03	1,25%

No trabalho de Westra *et al.* (2012), a utilização desse parâmetro pode ter sido essencial, devido ao grande número de estações empregadas e à grande quantidade de dados para amostragem, fazendo com que os autores tivessem que limitar a amostragem dos fragmentos a, no máximo, 10 dias. No caso da presente pesquisa, devido ao número limitado de estações e ao fato de que as mesmas apresentam em média apenas 6 anos de dados, há dificuldade de se encontrar mais de 10 dias dentro do limite de variação absoluta para amostragem. Frente ao exposto, o parâmetro (k) não foi utilizado no trabalho para nenhuma das durações, diminuindo assim uma variável no modelo.

Após o modelo calibrado, foram utilizadas as 1,000 séries para desagregação na etapa de avaliação final, considerando o limite da variação absoluta da precipitação de 5%, o número de estações vizinhas (S) igual a 10 e a não utilização do parâmetro número máximo de vizinhos (k = 10).

#### Durações de 180 minutos e 360 minutos

Os resultados da etapa de calibração para as durações de 180 e 360 minutos podem ser vistos na Tabela 5.10 e na Tabela 5.11. Observa-se por esses resultados que o modelo também apresentou valores bem próximos considerando o limite da variação absoluta da precipitação de 5% ou 10% e a variação do número de estações vizinhas entre 5, 10 e 15. Para essas duas



durações, o modelo apresentou resultados semelhantes e, por isso, as mesmas foram agrupadas neste subitem.

Como pode ser observado, há uma subestimação da variância e da média para quase todos os casos, mas mantendo uma variação abaixo de 7% para todas as métricas. Nesse caso, a utilização do limite de 10% trouxe resultados ligeiramente melhores, principalmente com a utilização de 15 vizinhos. Esse comportamento indica a necessidade de mais dados para a amostragem no caso dessas durações.

**Tabela 5.10 – Resultados das simulações de calibração – Duração 180 minutos**

<b>Duração de 180 min - Limite = 5% - Precipitação em (mm)</b>							
<b>Métricas</b>	<b>Observado</b>	<b>Nº. Vizinhos = 5</b>		<b>Nº. Vizinhos = 10</b>		<b>Nº. Vizinhos = 15</b>	
<b>Média dos máximos anuais</b>	51,23	52,56	2,61%	51,99	1,49%	51,76	1,05%
<b>Média</b>	4,35	4,16	-4,43%	4,18	-3,95%	4,19	-3,73%
<b>Variância</b>	50,26	48,76	-2,98%	49,03	-2,46%	49,05	-2,41%
<b>Assimetria</b>	3,20	3,41	6,60%	3,34	4,56%	3,30	3,30%
<b>Duração de 180 min - Limite = 10% - Precipitação em (mm)</b>							
<b>Métricas</b>	<b>Observado</b>	<b>Nº. Vizinhos = 5</b>		<b>Nº. Vizinhos = 10</b>		<b>Nº. Vizinhos = 15</b>	
<b>Média dos máximos anuais</b>	51,23	52,40	2,29%	52,41	2,30%	51,65	0,83%
<b>Média</b>	4,35	4,21	-3,09%	4,22	-2,88%	4,24	-2,61%
<b>Variância</b>	50,26	49,90	-0,71%	50,27	0,03%	49,94	-0,63%
<b>Assimetria</b>	3,20	3,35	4,72%	3,34	4,43%	3,26	1,98%

**Tabela 5.11 – Resultados das simulações de calibração – Duração 360 minutos**

<b>Duração de 360 min - Limite = 5% - Precipitação em (mm)</b>							
<b>Métricas</b>	<b>Observado</b>	<b>Nº. Vizinhos = 5</b>		<b>Nº. Vizinhos = 10</b>		<b>Nº. Vizinhos = 15</b>	
<b>Média dos máximos anuais</b>	60,10	61,47	2,28%	60,90	1,34%	60,42	0,53%
<b>Média</b>	6,09	5,77	-5,22%	5,80	-4,80%	5,79	-4,92%
<b>Variância</b>	86,61	85,58	-1,20%	86,72	0,12%	85,60	-1,17%
<b>Assimetria</b>	2,79	2,98	7,02%	2,92	4,84%	2,89	3,76%
<b>Duração de 360 min - Limite = 10% - Precipitação em (mm)</b>							
<b>Métricas</b>	<b>Observado</b>	<b>Nº. Vizinhos = 5</b>		<b>Nº. Vizinhos = 10</b>		<b>Nº. Vizinhos = 15</b>	
<b>Média dos máximos anuais</b>	60,10	61,70	2,66%	61,63	2,55%	60,44	0,57%
<b>Média</b>	6,09	5,86	-3,77%	5,86	-3,75%	5,85	-4,01%
<b>Variância</b>	86,61	88,10	1,72%	88,57	2,25%	86,91	0,34%
<b>Assimetria</b>	2,79	2,95	5,77%	2,93	4,93%	2,87	3,03%

Frente ao exposto, considerou-se o limite da variação absoluta da precipitação de 10% e o número de estações vizinhas igual a 15 para desagregação das 1.000 séries da etapa de avaliação final, tanto para a duração de 180 minutos, quanto para a de 360 minutos.

#### Duração de 720 minutos

Os resultados da etapa de calibração para a duração de 720 minutos podem ser vistos na Tabela 5.12. Observa-se, por esses resultados, que o modelo também apresentou valores bem próximos considerando o limite da variação absoluta da precipitação de 5% ou 10% e a variação do número de estações vizinhas entre 5, 10 e 15. Entretanto, considerando o limite da variação absoluta da precipitação de 5% e o número de estações vizinhas de 10, o modelo apresentou melhores resultados.

**Tabela 5.12 – Resultados das simulações de calibração – Duração 720 minutos**

<b>Duração de 720 min - Limite = 5% - Precipitação em (mm)</b>							
<b>Métricas</b>	<b>Observado</b>	<b>Nº. Vizinhos = 5</b>		<b>Nº. Vizinhos = 10</b>		<b>Nº. Vizinhos = 15</b>	
<b>Média dos máximos anuais</b>	71,31	70,23	-1,51%	71,24	-0,10%	73,54	3,13%
<b>Média</b>	8,53	8,23	-3,59%	8,16	-4,37%	8,18	-4,11%
<b>Variância</b>	147,51	146,47	-0,71%	148,76	0,85%	152,47	3,36%
<b>Assimetria</b>	2,52	2,52	0,12%	2,57	2,15%	2,64	4,98%
<b>Duração de 720 min - Limite = 10% - Precipitação em (mm)</b>							
<b>Métricas</b>	<b>Observado</b>	<b>Nº. Vizinhos = 5</b>		<b>Nº. Vizinhos = 10</b>		<b>Nº. Vizinhos = 15</b>	
<b>Média dos máximos anuais</b>	71,31	71,08	-0,31%	71,70	0,55%	73,86	3,57%
<b>Média</b>	8,53	8,28	-2,95%	8,20	-3,90%	8,23	-3,54%
<b>Variância</b>	147,51	149,42	1,30%	150,65	2,13%	154,31	4,61%
<b>Assimetria</b>	2,52	2,53	0,47%	2,57	2,19%	2,64	5,08%

Para essa duração o modelo apresentou uma subestimação da média, mantendo, contudo, uma variação de, no máximo, 5% para todas as métricas. Assim como observado para a duração de 60 minutos, a utilização de 5 e 15 vizinhos resultou em piora no desempenho do modelo. Esperava-se para a duração de 720 minutos o mesmo comportamento apresentado para as durações de 180 e 360 minutos. Visto que não há uma explicação clara do motivo do comportamento similar àquele apresentado pela duração de 60 minutos, a decisão de manter o modelo com 10 vizinhos foi puramente matemática, devido à maior eficiência das métricas.

Frente ao exposto, considerou-se o limite da variação absoluta da precipitação de 5% e o número de estações vizinhas igual a 10 para desagregação das 1.000 séries para etapa de avaliação final para a duração de 720 minutos.

#### 5.4 Validação e avaliação da eficiência do modelo

Foram desagregadas as 1.000 séries considerando os parâmetros de calibração escolhidos para cada uma das durações. As Figuras 5.8 a 5.11 apresentam os resultados obtidos para as métricas avaliadas, quais sejam, média dos máximos anuais, média, variância e assimetria das precipitações, considerando apenas os dias chuvosos, para as durações de 60, 180, 360 e 720 minutos, respectivamente. Em azul é apresentado, para cada uma das métricas, o valor obtido da série observada, ou seja, a série do pluviógrafo.

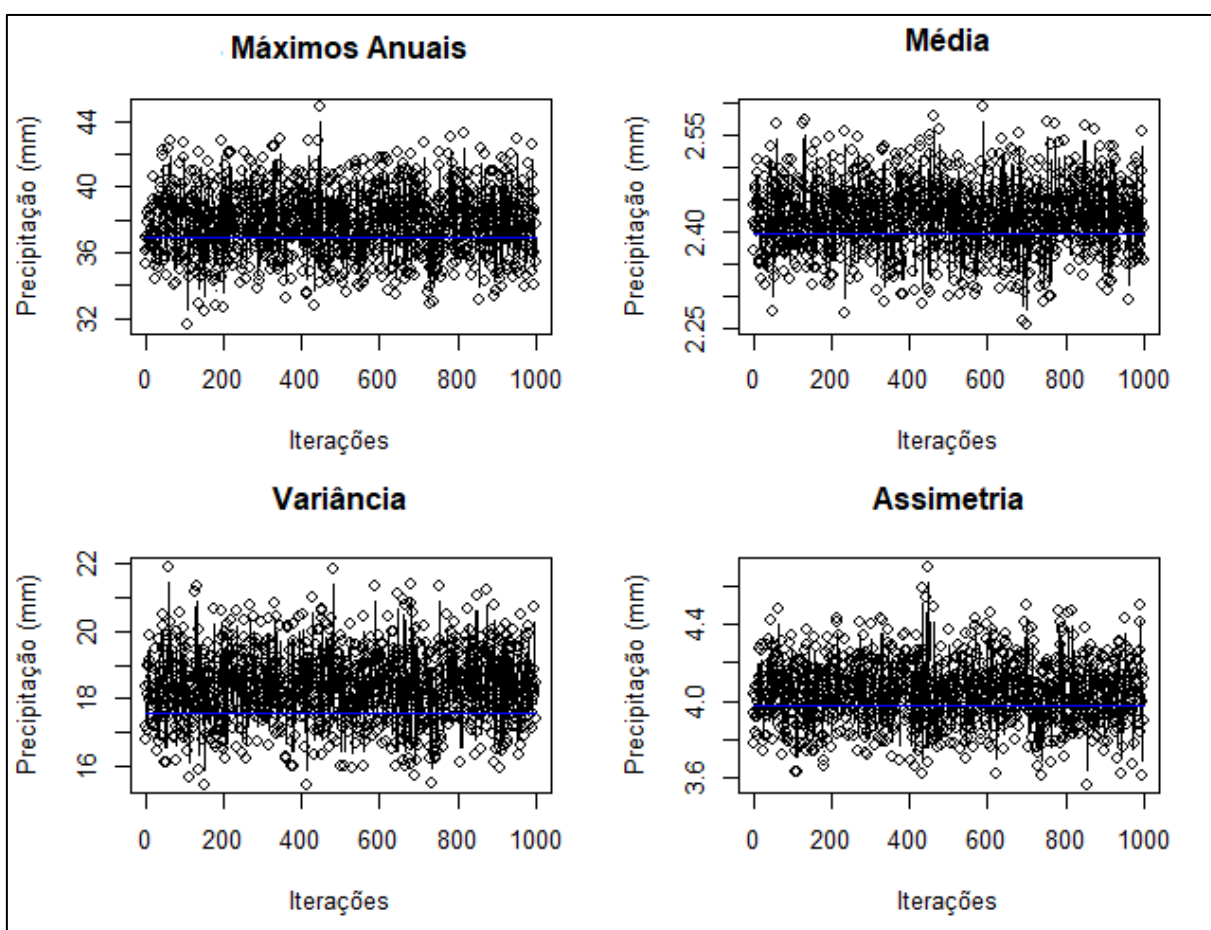
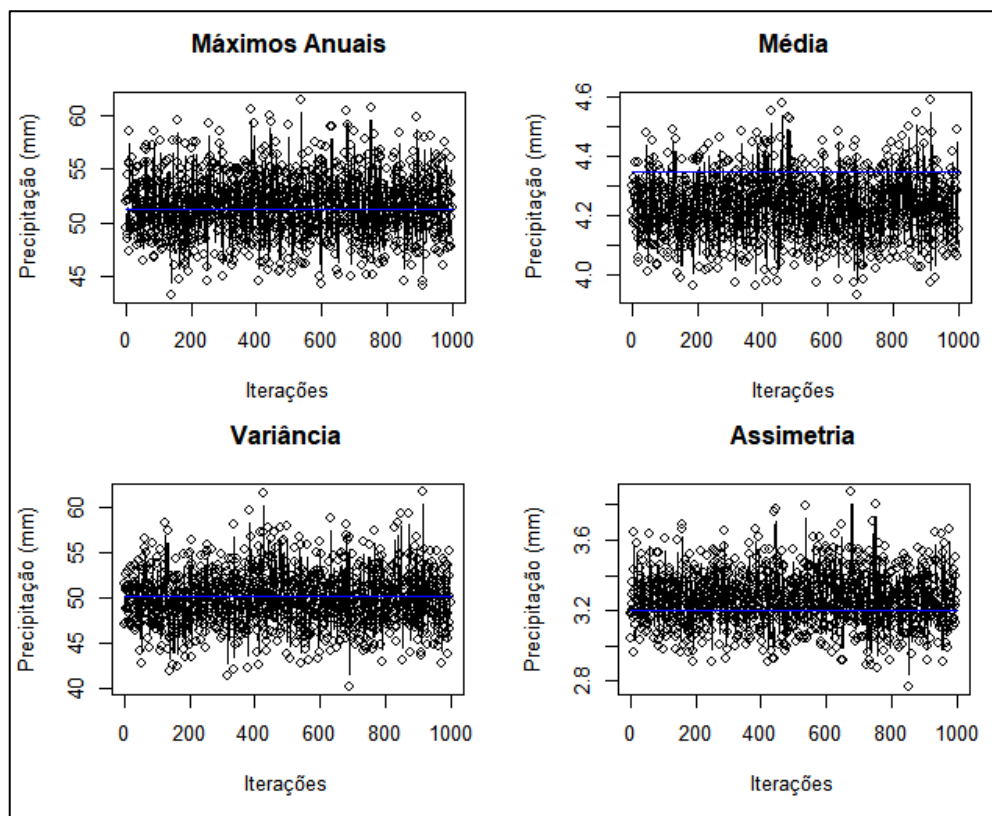
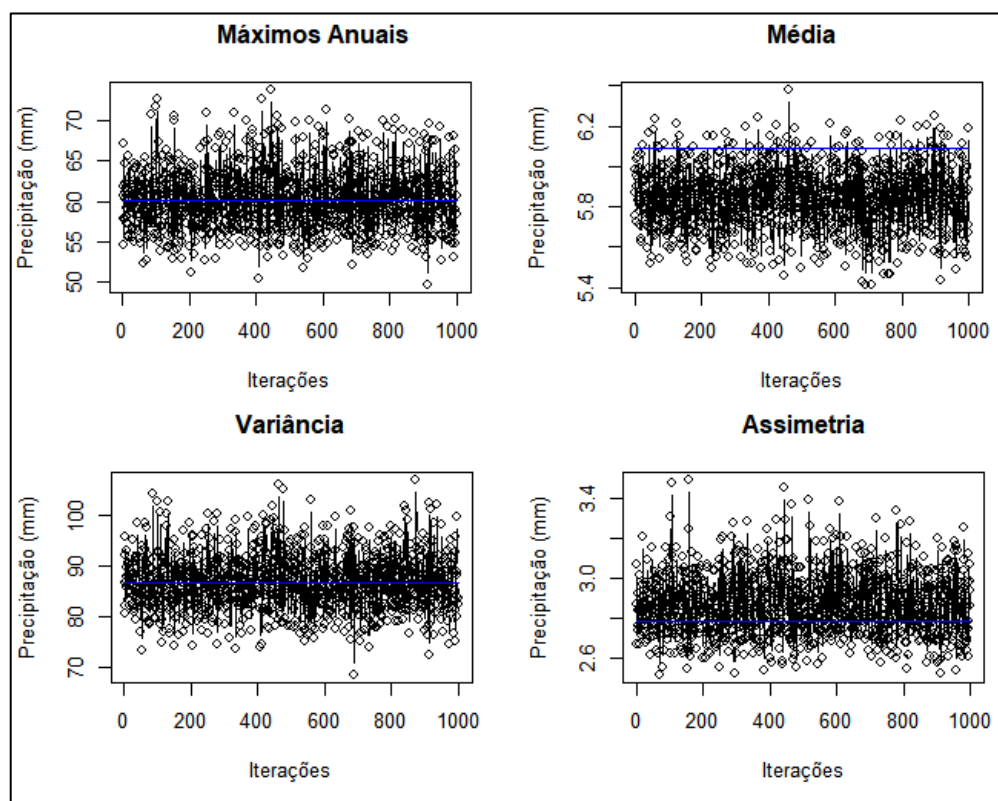


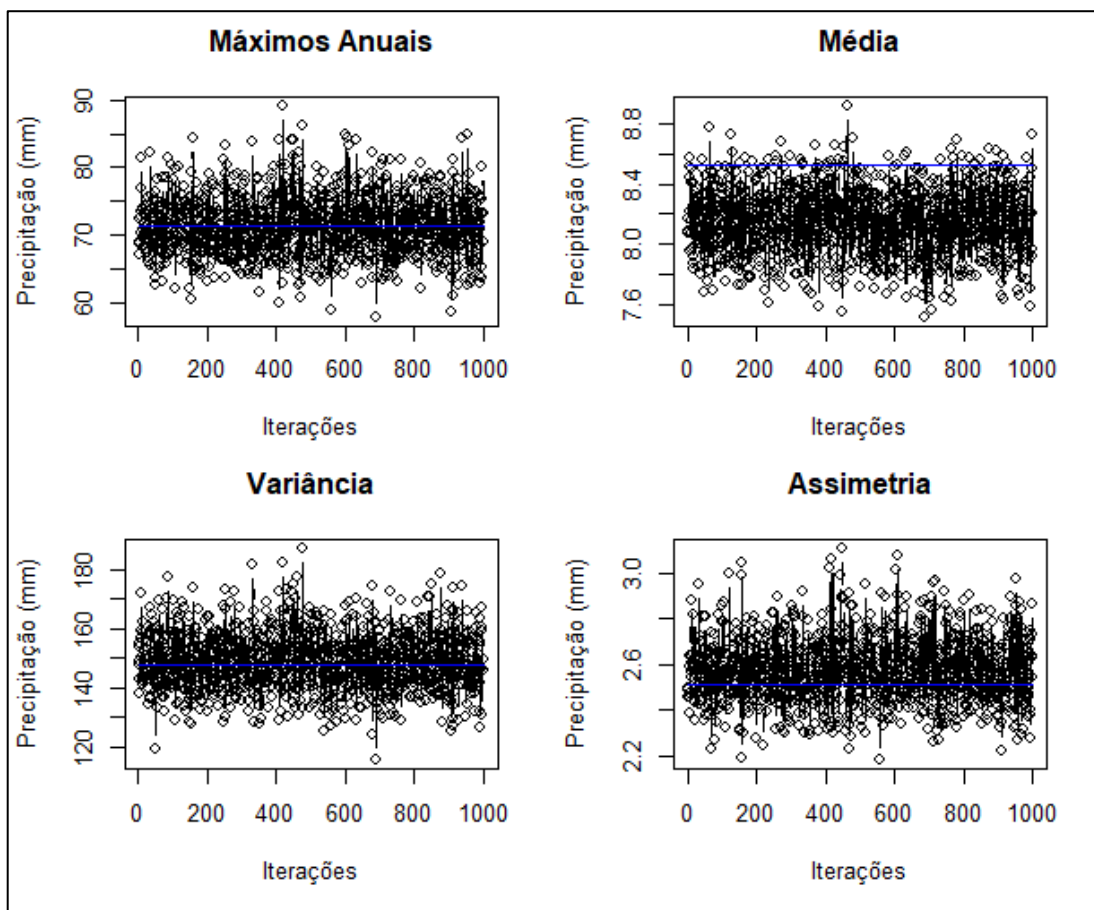
Figura 5.8 – Resultados das métricas obtidas para a duração de 60 min



**Figura 5.9** – Resultados das métricas obtidas para a duração de 180 min



**Figura 5.10** – Resultados das métricas obtidas para a duração de 360 min



**Figura 5.11** – Resultados das métricas obtidas para a duração de 720 min

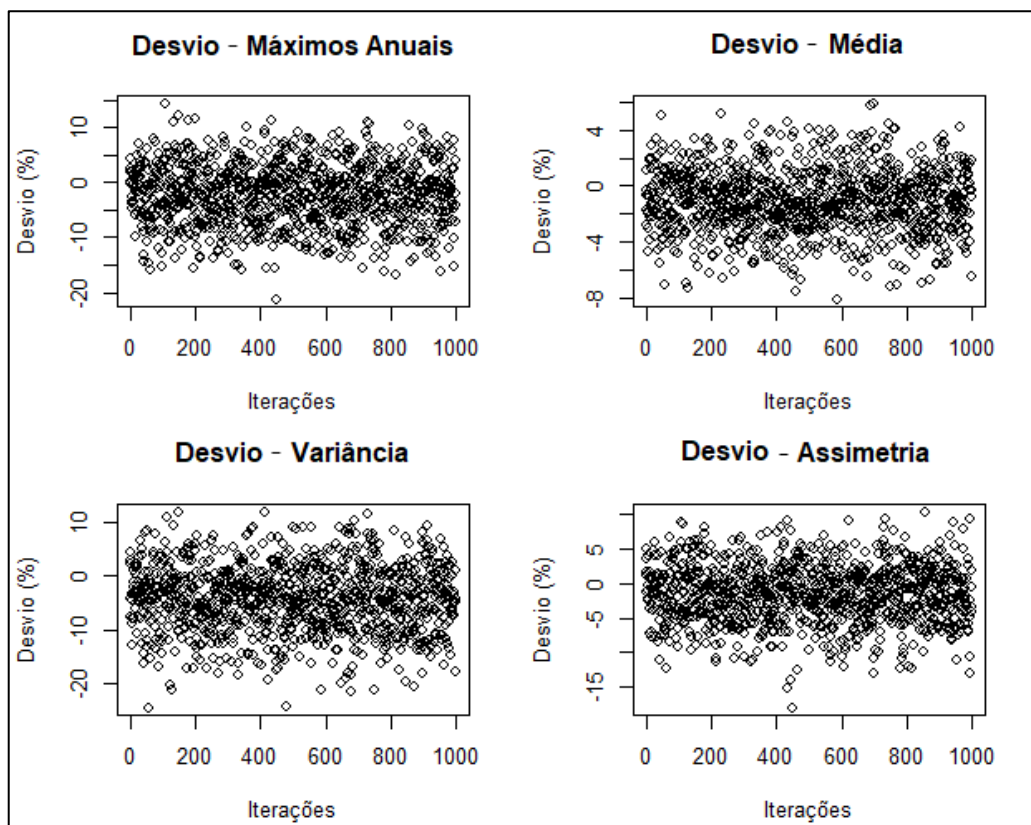
Observa-se que os valores simulados são próximos aos observados, o que é corroborado pela Tabela 5.13. Ressalta-se que, na referida tabela, são apresentados os valores observados, os valores simulados e o desvio entre os valores simulados e observados apresentado em porcentagem, com valores negativos indicando subestimação e positivos superestimação do valor observado.

Avaliando as figuras e a tabela mencionadas, observa-se que os resultados indicam uma tendência de superestimação para todas as métricas e durações, com exceção da média para a duração de 720 minutos, onde há uma tendência de subestimação de aproximadamente 4,5%.

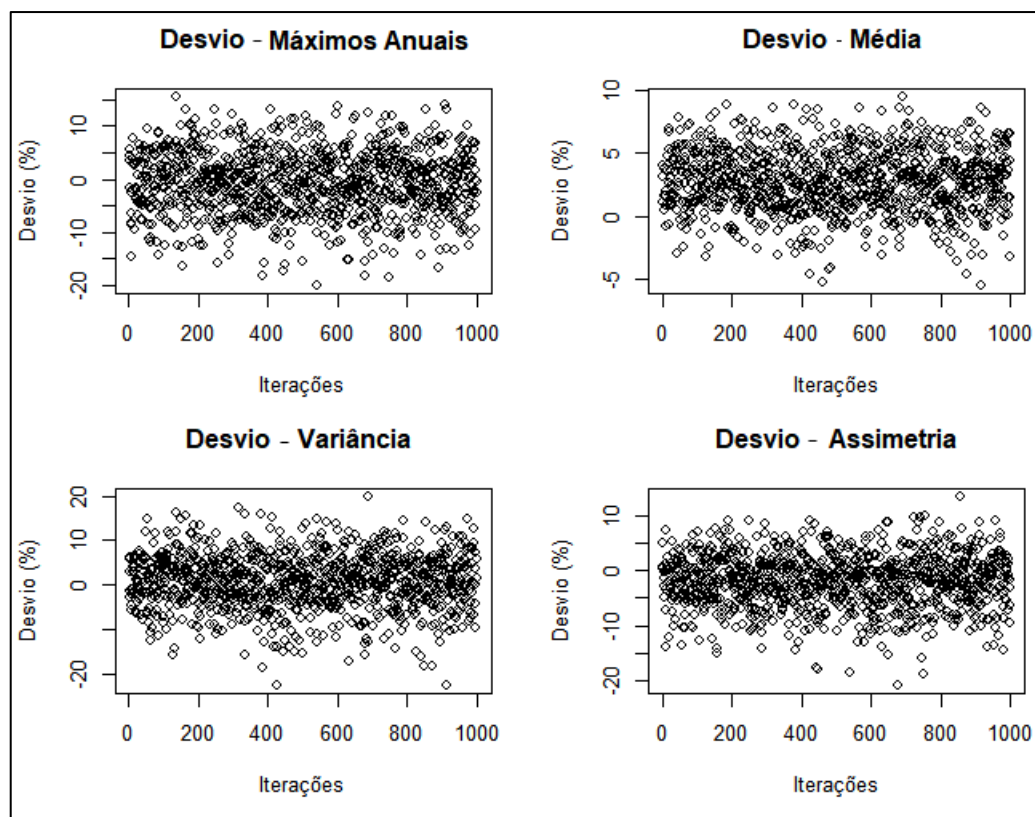
**Tabela 5.13** – Resultados das simulações com 1.000 iterações para todas as durações

<b>Nº Vizinhos = 10, Limite= 5%, N= 1000</b>						
<b>Durações</b>	<b>Duração de 60 minutos</b>			<b>Duração de 720 minutos</b>		
<b>Métricas</b>	<b>Observado</b>	<b>Simulado</b>		<b>Observado</b>	<b>Simulado</b>	
<b>Média dos máximos anuais</b>	36,98	37,81	2,26%	71,31	71,34	0,04%
<b>Média</b>	2,40	2,42	1,10%	8,53	8,15	-4,49%
<b>Variância</b>	17,58	18,35	4,38%	147,51	148,70	0,81%
<b>Assimetria</b>	3,98	4,04	1,59%	2,52	2,57	2,28%
<b>Nº Vizinhos = 15, Limite= 10%, N= 1000</b>						
<b>Durações</b>	<b>Duração de 180 minutos</b>			<b>Duração de 360 minutos</b>		
<b>Métricas</b>	<b>Observado</b>	<b>Simulado</b>		<b>Observado</b>	<b>Simulado</b>	
<b>Média dos máximos anuais</b>	51,23	51,60	0,73%	60,10	60,27	0,28%
<b>Média</b>	4,35	4,23	-2,69%	6,09	5,84	-4,09%
<b>Variância</b>	50,26	49,93	-0,67%	86,61	86,70	0,10%
<b>Assimetria</b>	3,20	3,26	2,05%	2,79	2,87	3,01%

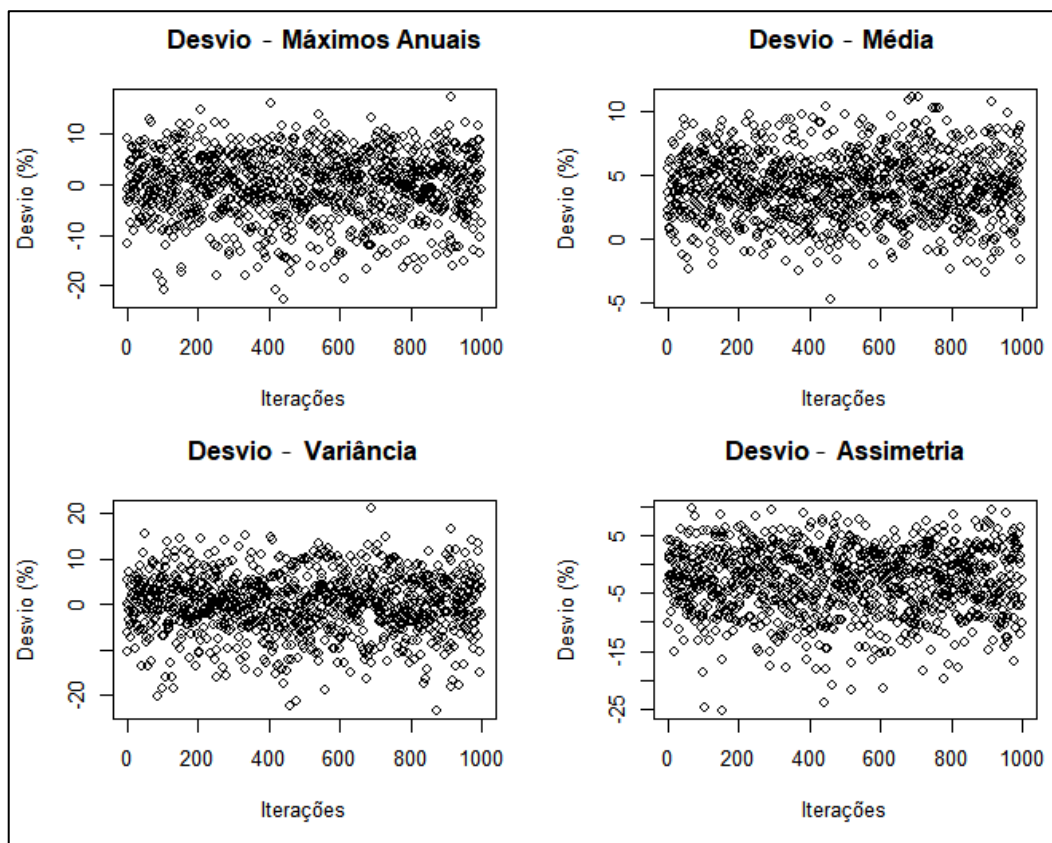
Já as Figuras 5.12 a 5.15 ilustram os resultados de desvio entre os valores simulados e os observados para todas as iterações, considerando os parâmetros de melhor eficiência, para as métricas média dos máximos anuais, média, variância e assimetria das precipitações para as durações de 60, 180, 360 e 720 min, respectivamente. Observa-se que grande parte dos resultados se encontram dentro de uma faixa de variação de  $\pm 10\%$ . Para a duração de 60 minutos, no caso da média e da assimetria, essa faixa é reduzida para  $\pm 5\%$ .



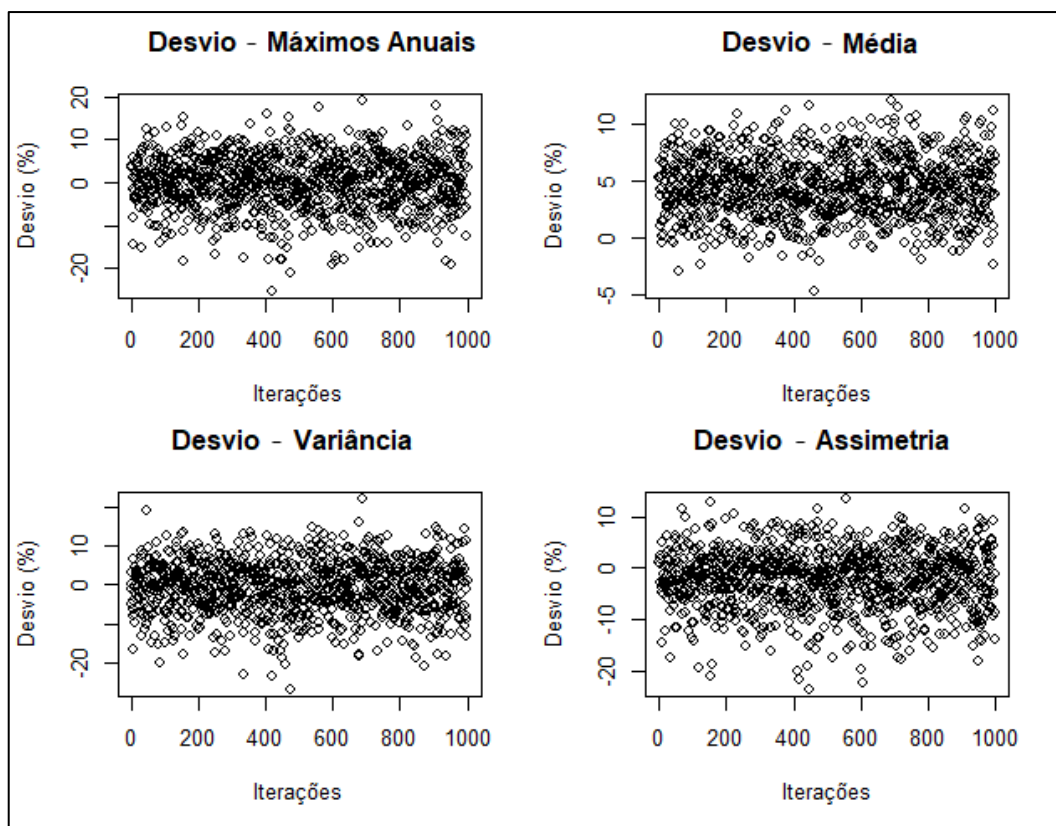
**Figura 5.12** – Resultados do desvio das métricas obtidas para a duração de 60 min



**Figura 5.13** – Resultados do desvio das métricas obtidas para a duração de 180 min



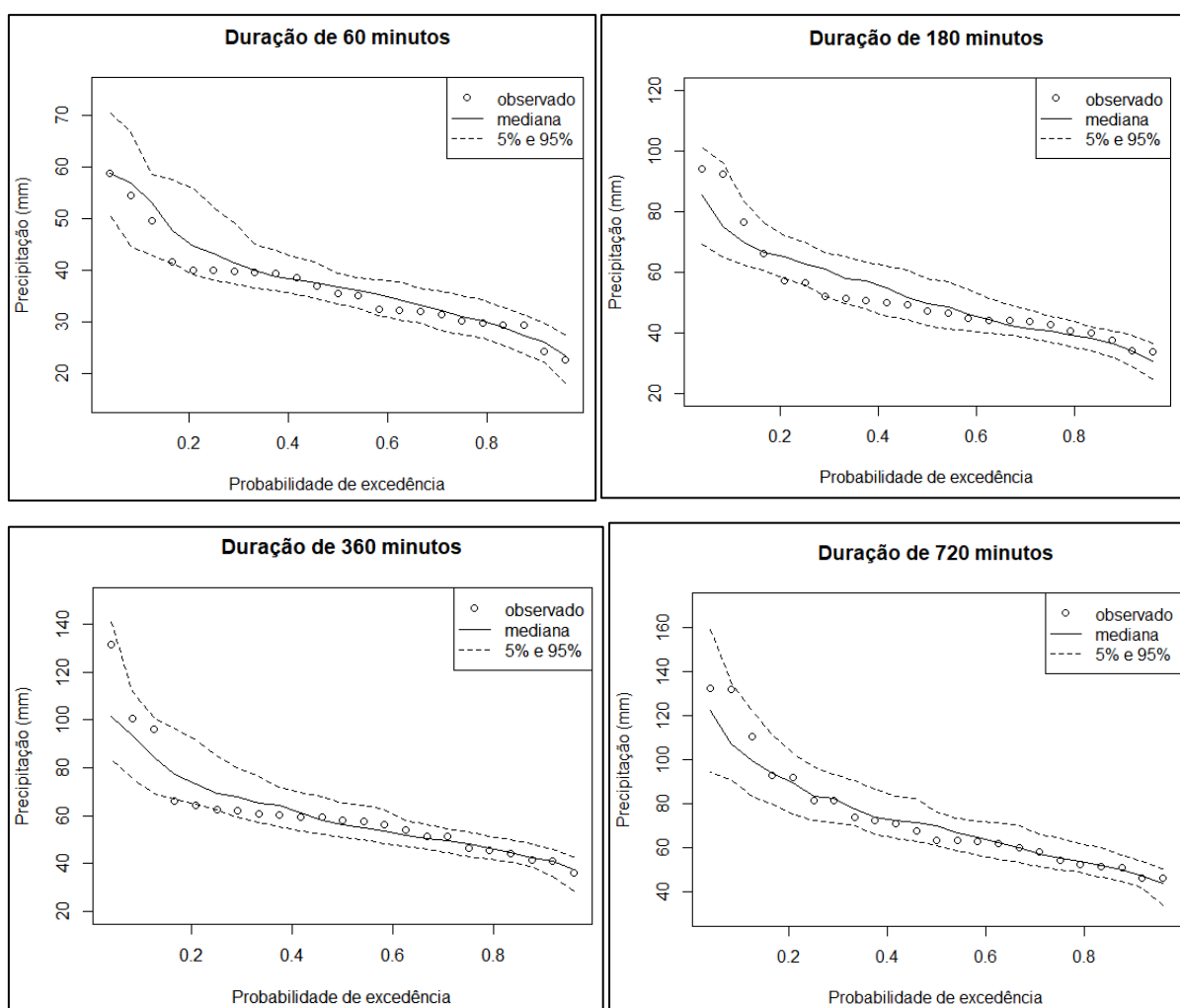
**Figura 5.14** – Resultados do desvio das métricas obtidas para a duração de 360 min



**Figura 5.15** – Resultados do desvio das métricas obtidas para a duração de 720 min



Para avaliação dos valores extremos, utilizou-se a equação de posição de plotagem de Weibull (Equação 5.1) para estimação da probabilidade empírica de excedência anual, a qual é expressa como uma fração entre 0 e 1. Essa equação permite a obtenção de probabilidades de excedência não enviesadas. A Figura 5.16 apresenta as precipitações máximas anuais para a duração de 60 minutos, plotadas em função das probabilidades empíricas de excedência anual, resultantes das 1.000 séries desagregadas pelo modelo na estação de Caeté. Os pontos representam os valores registrados pelo pluviógrafo, a linha contínua representa a mediana das 1.000 simulações e as linhas tracejadas indicam os percentis de 5% e 95% das séries simuladas, ou seja, o intervalo de confiança de 90%.



**Figura 5.16** – Curvas de quantis das precipitações máximas anuais para a estação de Caeté para todas as durações

Nota-se que os valores observados encontram-se dentro da faixa de variação do modelo, indicando que o mesmo consegue representar bem a variabilidade dos dados observados. Já no

trabalho de Westra *et al.* (2012), vários dos pontos observados caíam fora do intervalo dos percentis de 5 e 95, indicando uma subestimação da real variabilidade dos quantis. O melhor comportamento do modelo utilizado nesta pesquisa se deve à utilização do gerador diário, o qual agrega variabilidade às alturas de precipitação diária a serem desagregadas.

Outro ponto observado na Figura 5.16 é a tendência de superestimação da mediana. Li *et al.* (2018) encontraram uma tendência de subestimação para os eventos máximos anuais de menor probabilidade de excedência, e superestimação para os eventos de maior probabilidade de excedência. Esse comportamento também é observado, com exceção da duração de 60 minutos. Esse fato também pode ser explicado pela utilização do modelo diário, gerando valores mais extremos para a desagregação. Nota-se também, por esses resultados, um aumento da variabilidade para os eventos de baixa probabilidade de excedência e conseqüentemente de maior magnitude, assim como observado também por Westra *et al.* (2012).

De modo a comparar o modelo com uma análise de frequência tradicional, utilizou-se uma distribuição GEV para modelagem dos máximos anuais observados para duração de 60 minutos e 180 minutos com o mesmo intervalo de confiança (IC) de 90%. Os resultados dessa análise são apresentados na Tabela 5.14 e na Tabela 5.16. Utilizou-se o método da máxima verossimilhança para estimação dos parâmetros da GEV, cujos valores foram de:  $\kappa = 0,038$ ,  $\sigma = 6,955$  e  $\beta = 32,96$  para duração de 60 minutos; e  $\kappa = -0,032$ ,  $\sigma = 9,048$  e  $\beta = 44,22$  para duração de 180 minutos, sendo  $\kappa$  o parâmetro de forma,  $\sigma$  o parâmetro de escala e  $\beta$  o parâmetro de posição. Nota-se por essas tabelas que a ordem de grandeza dos intervalos para um mesmo período de retorno é similar. Os valores do modelo tendem a ficar mais próximos dos observados com o aumento do tempo de retorno para duração de 60 minutos. Já para a duração de 180 minutos, observa-se, em ambos (modelo e GEV), uma subestimação do valor observado, principalmente com o aumento do tempo de retorno, mas ainda encontram-se dentro do intervalo de confiança. Para as demais durações o comportamento é similar ao da duração de 180 minutos.

**Tabela 5.14** – Comparação do modelo com uma Distribuição GEV – Duração de 60 min

TR (anos)	Precipitação Observado (mm)	Precipitação GEV (mm)	Precipitação Modelo Mediana (mm)	IC - GEV (mm)	IC - Modelo (mm)
2	35,53	35,49	36,76	[32,56; 38,42]	[33,45; 39,49]
5	40,40	43,10	45,22	[39,04; 47,16]	[39,58; 56,20]
10	52,05	47,96	55,13	[42,51; 53,42]	[43,92; 62,80]
24	58,84	53,92	58,79	[45,51; 62,32]	[50,92; 70,63]

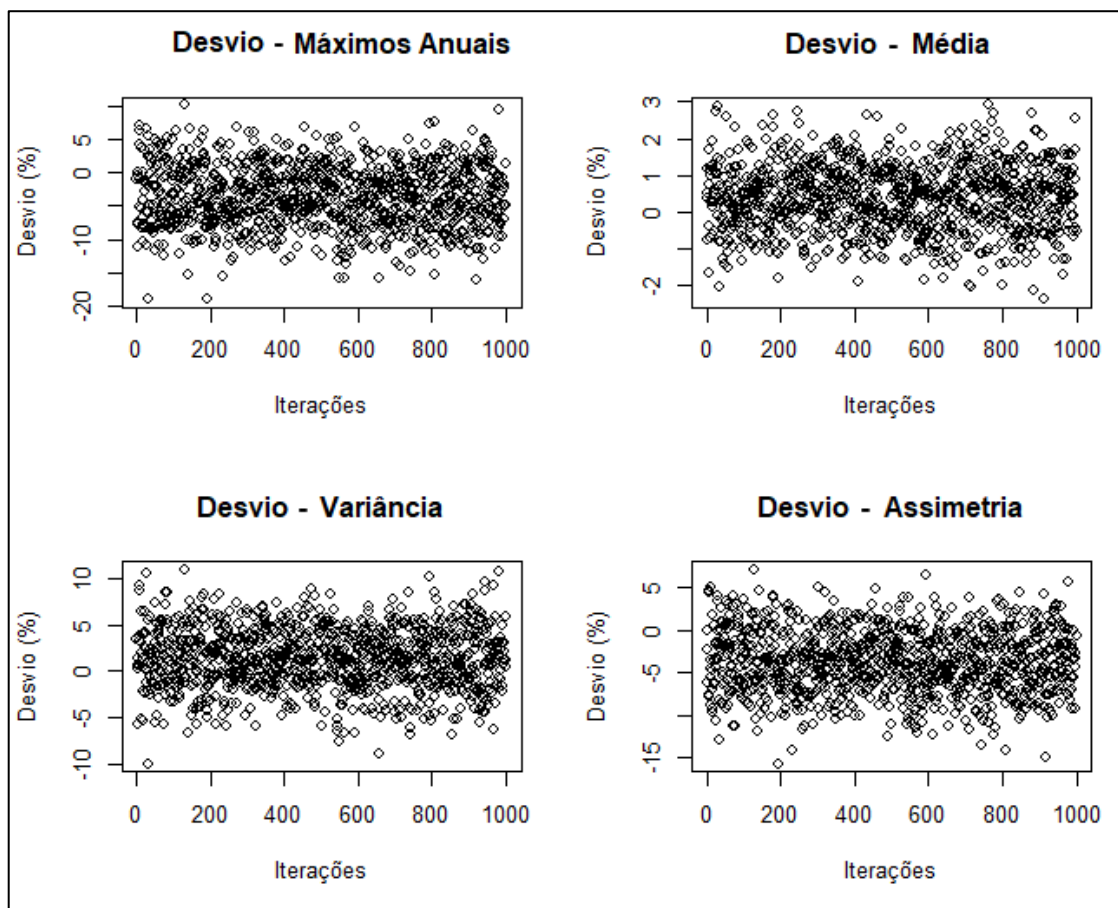
**Tabela 5.15** – Comparação do modelo com uma Distribuição GEV – Duração de 180 min

TR (anos)	Precipitação Observada (mm)	Precipitação GEV (mm)	Precipitação Modelo Mediana (mm)	IC - GEV (mm)	IC - Modelo (mm)
2	47,35	47,68	49,54	[43,47; 51,89]	[42,37; 58,06]
5	58,55	60,45	65,35	[52,47; 68,43]	[58,49; 72,79]
10	84,55	70,95	72,62	[57,54; 84,37]	[63,70; 89,86]
24	94,13	87,12	85,55	[61,17; 113,1]	[69,27; 101,44]

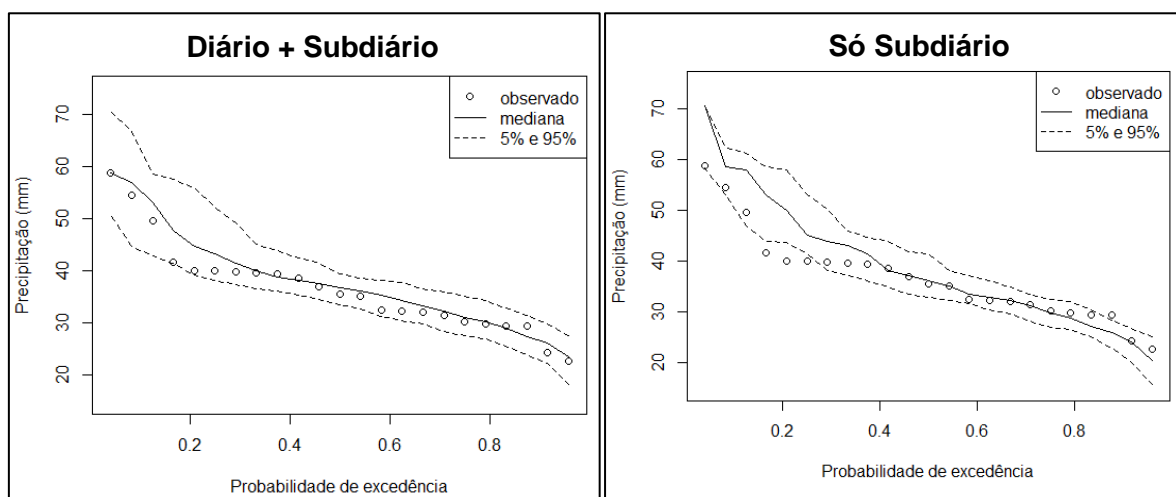
Visando confirmar o comportamento observado pelo modelo desta pesquisa frente a outras abordagens, como em Westra *et al.* (2012), Li *et al.* (2018), Pui *et al.* (2012) e Lu & Qin (2014), que consistem apenas no gerador subdiário, realizou-se uma simulação utilizando esse gerador para comparação. Os resultados dessa análise encontram-se na Tabela 5.16 e nas Figuras 5.18 e 5.19.

**Tabela 5.16** – Resultados das simulações – Diferenças entre modelos – Duração 60 minutos

Nº. Vizinhos = 10, Limite = 5%, N = 1000					
Métricas	Observado	Diário +Subdiário		Só subdiário	
Média dos máximos anuais	36,98	37,81	2,26%	38,32	3,62%
Média	2,40	2,42	1,10%	2,39	-0,37%
Variância	17,58	18,35	4,38%	17,35	-1,34%
Assimetria	3,98	4,04	1,59%	4,13	3,75%



**Figura 5.17** – Resultados do desvio das métricas obtidas para a duração de 60 min do modelo unicamente subdiário

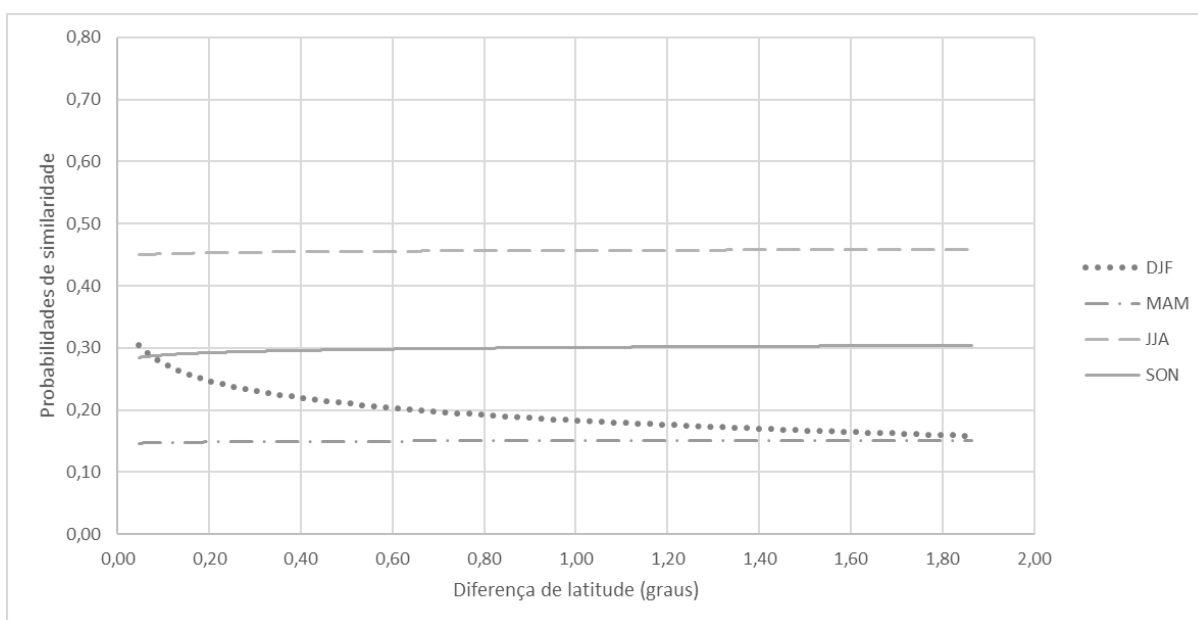


**Figura 5.18** – Curvas de quantis das precipitações máximas para a duração de 60 minutos

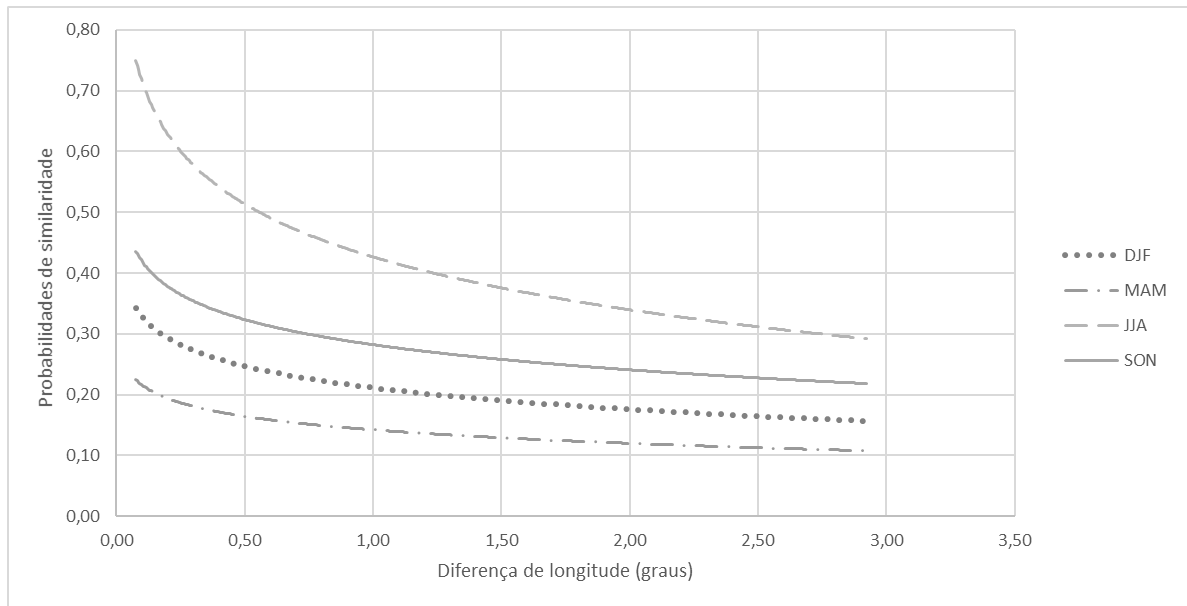
Conforme indicado anteriormente, observa-se, por esses resultados, que o modelo acoplado se comporta melhor que o modelo unicamente subdiário, estimando de maneira mais acurada as estatísticas dos máximos anuais e da assimetria. Analisando o apresentado pelas figuras

anteriormente citadas, pode-se concluir que a adição do gerador diário agrega variabilidade às alturas de chuva a serem desagregadas, aumentando os intervalos de confiança das simulações. Nesse caso, isso é uma vantagem para simulação de extremos, visto que vários pontos observados dos máximos anuais tendem a cair fora do intervalo nos modelos unicamente subdiários, tal como discutido por Westra *et al.* (2012) e apresentado pela Figura 5.18. Desse modo, o modelo apresenta uma melhor capacidade de reprodução de extremos em escala subdiária frente a outras abordagens.

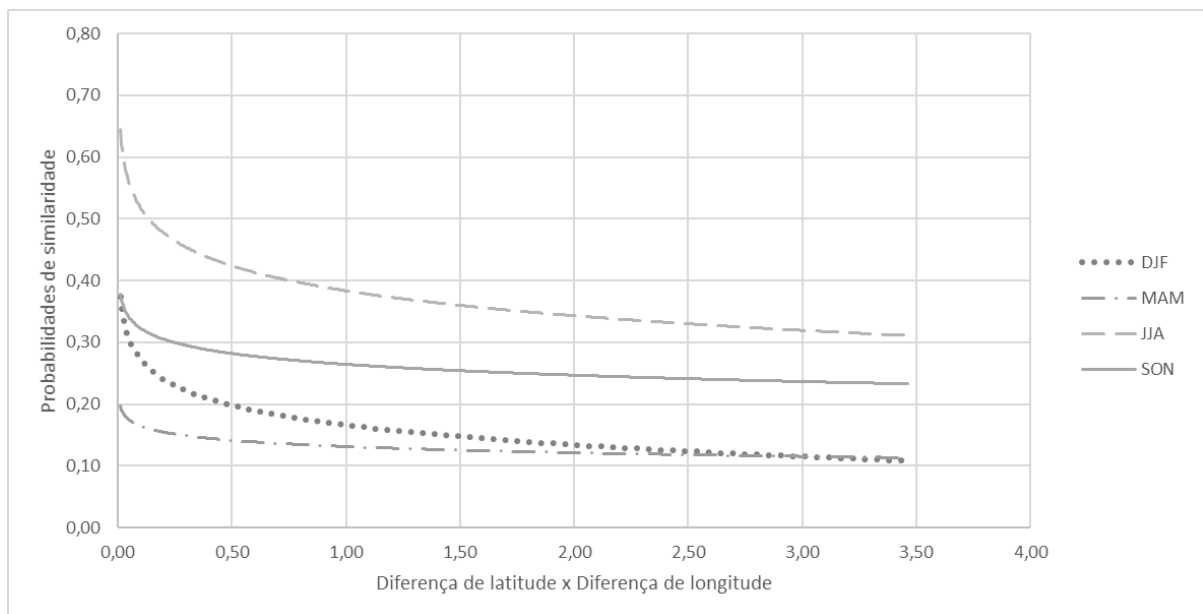
Outro aspecto do modelo também avaliado neste trabalho é a identificação das principais características para preservação da distribuição conjunta, ou seja, se alguma das características hidrológicas ou fisiográficas têm um maior impacto que as demais no modelo. Para responder essa questão, foram elaboradas as Figuras 5.19 a 5.22.



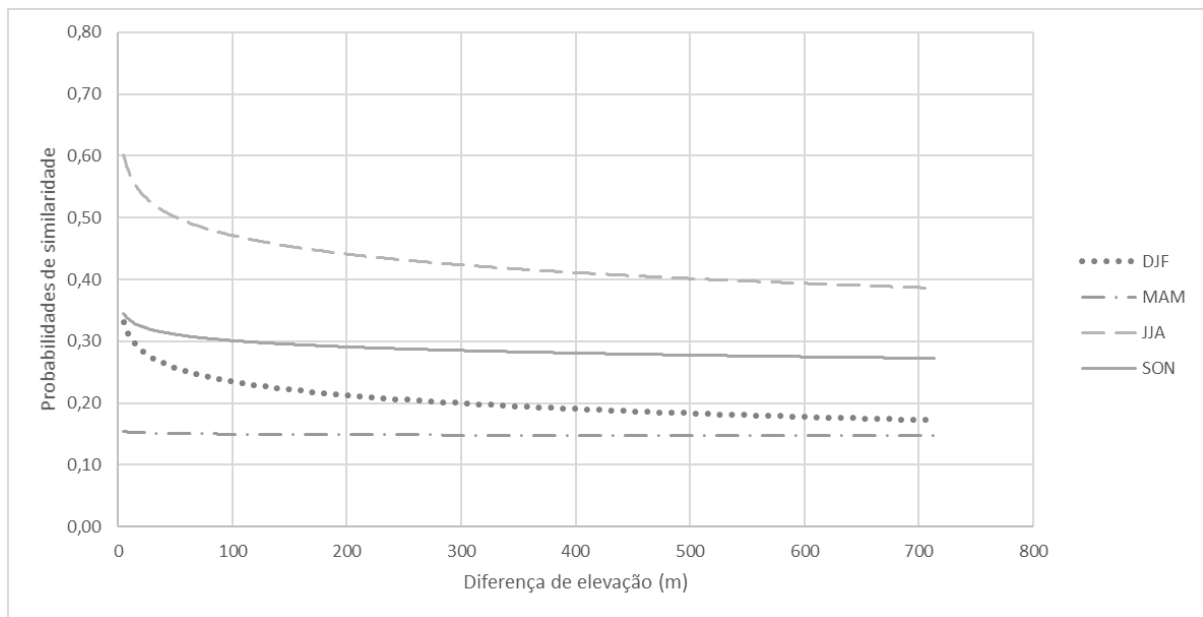
**Figura 5.19** – Probabilidade de similaridade em função da latitude para cada estação do ano



**Figura 5.20** – Probabilidade de similaridade em função da longitude para cada estação do ano



**Figura 5.21** – Probabilidade de similaridade em função da multiplicação da diferença de longitude e latitude para cada estação do ano



**Figura 5.22** – Probabilidade de similaridade em função da elevação para cada estação do ano

Pode-se notar que, para as estações mais chuvosas, as probabilidades de similaridade são menores, o que pode ser explicado pelo maior número de dias chuvosos, dificultando a obtenção de correlações entre as estações. Observa-se que, dentre as características fisiográficas, a longitude é a que possui maior influência nas probabilidades, já que existe uma diminuição das mesmas com o aumento da diferença de longitude. Esse fato também é corroborado pelos coeficientes do modelo de regressão, para os quais os maiores valores são referentes à característica fisiográfica longitude. Além disso, ressalta-se a questão da escala deste trabalho, que é reduzida a apenas três sub-bacias, resultando em diferenças das características fisiográficas menos pronunciadas, quando comparado ao trabalho de Westra *et al.* (2012).

Outro ponto interessante observado é que, para as características fisiográficas latitude e elevação, não há um impacto significativo com o aumento da diferença entre a estação de interesse e as demais estações. Isso indica que as características hidrológicas têm um impacto de maior magnitude, controlando o comportamento das probabilidades. No caso da latitude observa-se que não há influência, indicando que, mesmo estando mais distantes, as estações possuem similaridades hidrológicas que compensam essa grande diferença.

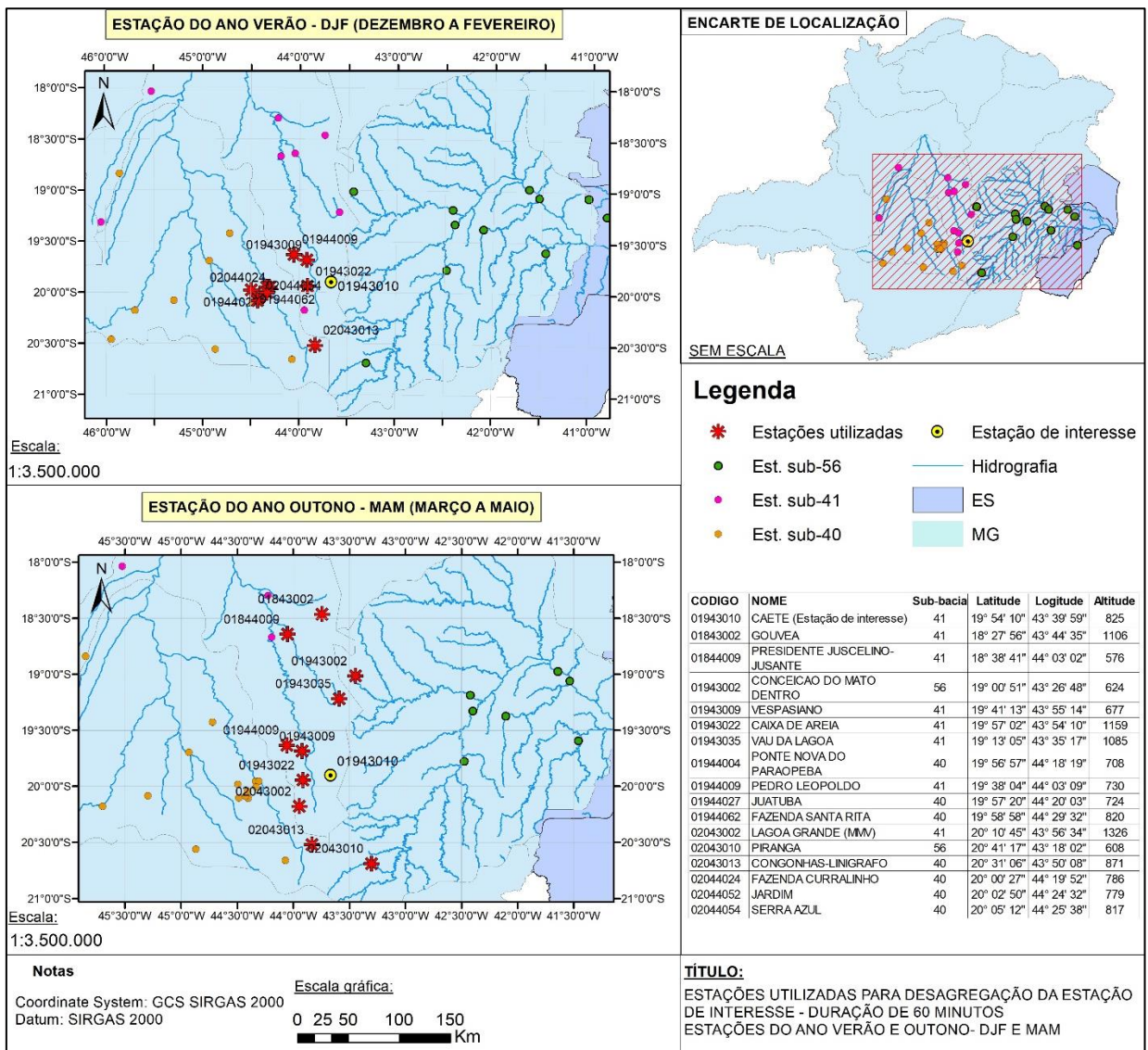
Para o atributo multiplicação da diferença de longitude e latitude, nota-se, que como a longitude tem um maior impacto que a latitude, o comportamento desse gráfico é similar ao gráfico para longitude. No caso da elevação há uma tendência de redução da similaridade, estabilizando para

diferenças maiores que 400 m, o que indica que, com diferenças de elevações superiores a esse valor, é mais difícil encontrar similaridade, o que é plausível considerando aspectos climáticos.

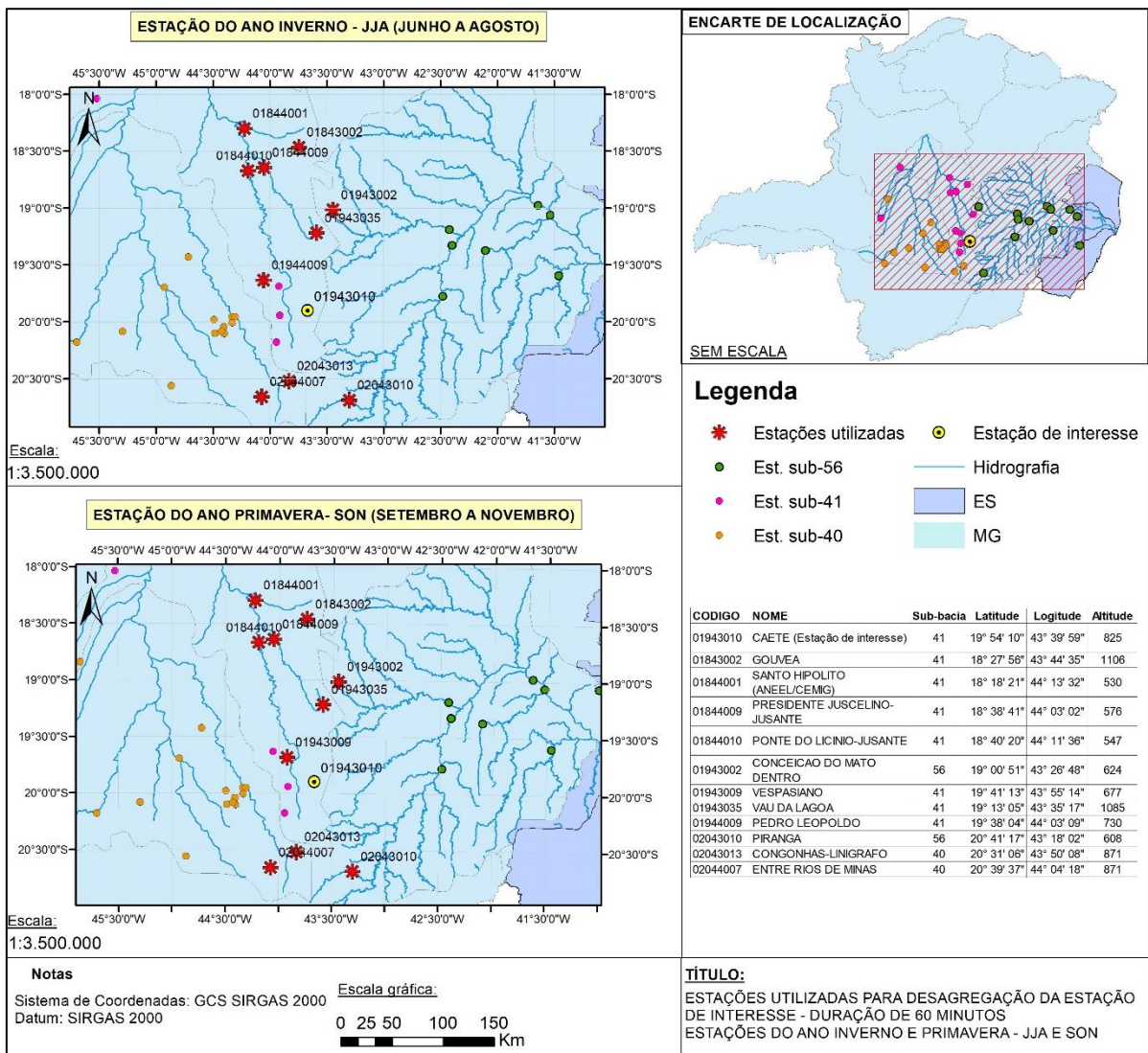
Já no caso das características hidrológicas, observa-se claramente um maior impacto do atributo intensidade máxima pelos altos valores de probabilidade resultantes da definição da similaridade entre as estações, conforme discutido anteriormente.

Por fim, de modo a visualizar as tendências encontradas nas probabilidades de similaridade no espaço, foram produzidas as Figuras 5.23 e 5.24. Nelas estão indicados os postos utilizados pelo modelo de desagregação, para cada uma das estações do ano, para a duração de 60 minutos. Para as demais durações, ver Figuras I.1 a I.6 localizadas no Apêndice I. Analisando a mesma estação do ano para todas as durações, observa-se que o modelo possui um comportamento similar. Para a estação do ano verão (DJF), o modelo seleciona como as estações mais similares aquelas mais próximas à estação de interesse. Para a estação do ano outono (MAM), o modelo já começa a selecionar estações mais distantes geograficamente, encontrando maiores similaridades, mas ainda selecionado várias estações próximas a estação de interesse. O comportamento do modelo para as estações do ano inverno (JJA) e primavera (SON) é similar, passando a identificar maiores similaridades em estações mais distantes. O contrário do observado para outras estações do ano, que encontram maiores similaridades em estações mais próximas à estação de interesse. Esse comportamento sugere que nas estações do ano menos chuvosas (JJA e SON) os atributos fisiográficos têm menor impacto, já que, com menos dias chuvosos, é mais fácil encontrar similaridades por meio dos atributos hidrológicos, corroborando o observado nas Figuras 5.19 a 5.22.





**Figura 5.23** – Estações utilizadas para desagregação – Duração 60 minutos – Estações do ano Verão e Outono



**Figura 5.24** – Estações utilizadas para desagregação – Duração 60 minutos – Estações do ano Inverno e Primavera

## 6 CONCLUSÕES

O desenvolvimento do presente trabalho buscou apresentar uma alternativa à indisponibilidade de dados de precipitação com alta resolução temporal, visto que amostras longas e contínuas de precipitação subdiária são, via de regra, difíceis de ser encontradas. Além disso, buscou incluir uma abordagem para simulação de extremos nessa escala temporal, adicionando ao modelo a capacidade de simular de maneira apropriada tanto os eventos regulares quanto aqueles mais extremos.

Para geração das chuvas diárias a serem desagregadas foi utilizado um gerador estocástico diário, composto por um modelo misto de simulação, que consiste na utilização da abordagem não paramétrica de reamostragem para os eventos regulares, da abordagem paramétrica para os eventos extremos e de um modelo multiestados para permitir a modelagem das chuvas de naturezas físicas distintas. Na abordagem para os eventos extremos, foi utilizada uma distribuição de probabilidade superiormente limitada, a saber, a distribuição Lognormal de 4 parâmetros (LN4), com base na hipótese de que a síntese física dos processos de formação de tormentas em condições extremas impõe um limite superior finito a essa variável. Mesmo sendo esse um tema controverso, alguns estudos fornecem evidências de que as precipitações são limitadas superiormente e, sendo assim, essa variável deve ser modelada sob tal perspectiva.

A definição de um limite superior foi realizada empregando uma estimativa da Precipitação Máxima Provável (PMP), porém de forma probabilística, incorporando a variabilidade e incertezas de estimação dessa grandeza. Tal estimativa foi incluída na estrutura do gerador por meio da especificação de uma distribuição *a priori* para o limite superior, na qual a PMP constitui um estimador para a precipitação máxima, e sua estimativa é associada a uma probabilidade de superação que certamente influenciará a estimativa do referido limite. Para os demais parâmetros da distribuição LN4, a ausência de relação direta com o fenômeno físico faz com que sejam empregadas distribuições *a priori* não informativas. Utilizando uma amostra e a função de verossimilhança específica, foram obtidas as estimativas pontuais dos parâmetros da distribuição LN4. De posse das estimativas, a referida distribuição foi empregada para a simulação das alturas de chuva no algoritmo do gerador estocástico de precipitação diária.

Em relação ao gerador diário, as principais conclusões obtidas foram:

- O limiar entre os eventos regulares e extremos foi definido como 100 mm, visto que esse se adaptou melhor ao conjunto de máximos anuais observados e que, em virtude das características climáticas e de sazonalidade da região de estudo, a probabilidade de ocorrência de um máximo anual fora do intervalo entre novembro e fevereiro é muito baixa. Assim, para preservar o comportamento da cauda superior das precipitações diárias e diminuir impactos para maiores durações, como nas médias mensais e anual, definiu-se pela escolha do limiar de 100 mm para o gerador estocástico diário;
- O modelo utilizado reproduz de maneira adequada as estatísticas diárias, para todos os meses do ano, confirmando a escolha adequada do limiar de 100 mm e do tamanho da janela de amostragem de 14 dias. Entretanto, um único limiar para todos os meses do ano dificulta a calibração do modelo;
- Com relação ao número médio de dias chuvosos, observou-se um viés de subestimação, mais evidente nos meses da estação chuvosa (novembro-março); e
- O gerador foi capaz de reproduzir de forma apropriada as médias e variância mensais, apresentando uma tendência de superestimação, o que constitui uma grande vantagem do modelo proposto em relação às alternativas estritamente paramétricas.

Após a geração de 1.000 séries com o modelo diário, as mesmas foram desagregadas utilizando o modelo subdiário, empregando uma abordagem de similaridade regional. Essa abordagem permite que os fragmentos de precipitação subdiária sejam aleatoriamente amostrados de pluviógrafos nas proximidades, condicionados à altura de chuva diária no local de interesse. A identificação das estações de maior similaridade foi realizada por meio de atributos hidrológicos como a intensidade máxima e o tempo da intensidade máxima, e atributos fisiográficos como a diferença absoluta entre elevações, latitudes, longitudes, e a multiplicação das diferenças de latitude e longitude. Após a identificação das probabilidades de similaridade à estação de interesse, o modelo passou por uma etapa de calibração, identificando as estações utilizadas pelo algoritmo de desagregação e o limite da variação da precipitação diária, para o qual apenas os fragmentos dentro dessa faixa de variação seriam amostrados.

Empregando o procedimento apresentado para cada uma das durações escolhidas (60, 180, 360 e 720 minutos), foram desagregadas as 1.000 séries de precipitação diária obtidas com o gerador

estocástico diário. Para validação do gerador, utilizou-se da estação de Caeté, a única disponibilizada com um maior número de dados subdiários (23 anos de dados).

Em relação ao gerador subdiário, as principais conclusões obtidas foram:

- Na identificação da similaridade, o atributo hidrológico intensidade máxima funcionou apropriadamente. Em relação ao atributo tempo da intensidade máxima, o comportamento apresentado indicou uma tendência de diminuição da similaridade com o aumento da duração. Já para o atributo fração de zeros, o comportamento apresentado não indicou similaridade para nenhuma das durações e estações do ano, com exceção para a estação do ano inverno (junho, julho e agosto), dificultando sua utilização neste trabalho;
- As probabilidades das estações mais similares para o atributo intensidade máxima se situaram entre 0,6 a 0,7, que são similares ao encontrados na literatura, com exceção das estações mais chuvosas, nas quais essa probabilidade é menor. Já para o atributo tempo da intensidade máxima as probabilidades de similaridade foram baixas;
- A característica fisiográfica de maior impacto para identificação da similaridade foi a longitude, confirmado pelos maiores coeficientes no modelo de regressão logística, o que parece razoável considerando a escala do trabalho;
- O número de estações vizinhas para amostragem dos fragmentos foi de 10 para as durações de 60 e 720 minutos e 15 para durações de 180 e 360 minutos, demonstrando que é importante a calibração do número de vizinhos escolhidos para a simulação. Já o limite da variação absoluta da precipitação foi de 5 % para as durações de 60 e 720 minutos e 10 % para durações de 180 e 360 minutos. Tal característica, no entanto, possui menor impacto do que o parâmetro (S);
- Observou-se que a utilização do número máximo de vizinhos (k) igual a 10 não possui impacto perceptível no modelo, visto o número limitado de estações e que há dificuldade de se encontrar mais de 10 dias dentro do limite de variação absoluta para amostragem;
- A avaliação dos geradores por meio das estatísticas subdiárias indicou que o modelo funcionou apropriadamente quando comparado com a estação de validação, qual seja,

o posto de Caeté. Entretanto, a indisponibilidade de mais estações para validação limita uma avaliação do total potencial do gerador;

- Os valores observados para os máximos anuais para cada duração encontram-se dentro da faixa de variação do modelo, indicando que o mesmo consegue representar bem a variabilidade dos dados observados, devido à utilização do gerador diário, o qual agrega variabilidade às alturas de chuva a serem desagregadas;
- Quando comparado a uma análise de frequência tradicional, utilizando uma distribuição GEV para modelagem dos máximos anuais observados, notou-se que a ordem de grandeza dos intervalos de confiança obtidos com o modelo proposto e com a GEV para um mesmo período de retorno é a mesma; e
- Quando comparado aos modelos unicamente subdiários na literatura, os resultados encontrados também indicam a melhor eficiência do modelo, conseguindo contornar o problema de subestimação da variância e simulação de eventos extremos, aspecto esse que indica que a utilização de um gerador diário acoplado a um gerador subdiário é uma boa alternativa.

## 7 RECOMENDAÇÕES

As conclusões apresentadas permitem afirmar que, de maneira geral, os objetivos da presente pesquisa foram alcançados. Contudo, alguns aspectos podem ser melhorados e pontos não explorados nas aplicações podem ser abordados em desenvolvimentos futuros. As principais recomendações nesse sentido são:

- Utilização de limiares em escala mensal ou diária, visando melhor calibrar o modelo diário em relação a simulação de extremos;
- A inclusão de um maior número de estações pluviográficas no modelo, visando a obtenção de uma gama maior de dados para identificação de similaridade e possível amostragem;
- A utilização de um maior número de estações pluviográficas para validação, com diferentes características hidrológicas e fisiográficas, buscando avaliar todo o potencial do gerador utilizado neste trabalho;
- A avaliação da influência da duração sobre as estatísticas do modelo. Além do emprego do modelo para durações inferiores, buscando avaliar o comportamento do mesmo para resoluções temporais ainda mais altas, como 5, 10, 15 e 30 minutos; e
- A busca por alternativas de programação para redução do alto custo computacional, favorecendo o emprego de um maior número de simulações e um número maior de estações para desagregação.

## REFERÊNCIAS

- BOTERO, B. A. Estimación de crecidas e alto período de retorno mediante funciones de distribución com limite superior e informacío no sistemática. 223 p. Tese (Doutorado em Engenharia Hidráulica y Medio Ambiente)–Departamento de Ingeniería Hidráulica y Medio Ambiente da Universidad Politécnica de Valencia, Espanha, 2006.
- BOTERO, B. A.; FRANCÉS, F. Estimation of high return period flood quantiles using additional non-systematic information with upper bounded statistical models. *Hydrol. Earth Syst. Sci.*, 14, 2617-2628, <https://doi.org/10.5194/hess-14-2617-2010>, 2010.
- BOUGHTON, W.C. A daily rainfall generating model for water yield and flood studies. CRC for Catchment Hydrology. ISBN: 9781876006495. 21p, Melbourne, 1999.
- BOUGUILA, N.; ZIOU, D.; MONGA E. Practical Bayesian estimation of a finite beta mixture through gibbs sampling and its applications. *Statistics and Computing*, 16: 215–225. doi:10.1007/s11222-006-8451-7, 2006.
- BUIHAND, T.A. Some remarks on the use of daily rainfall models. *Journal of Hydrology*, 36, pp. 295-308, 1978.
- BUIHAND, T. A.; BRANDSMA, T. Multisite simulation of daily precipitation and temperature in the Rhine Basin by nearest-neighbor resampling. *Water Resour. Res.*, 37(11), 2761–2776, 2001.
- CHAPMAN, T.G. Stochastic models for daily rainfall. *Water Down Under 94*. Institution of Engineers Australia 3: 7-12, 1994.
- CHAPMAN, T.G. Stochastic models for daily rainfall in the Western Pacific. *Mathematics and Computers in Simulation*, 43, pp. 351- 358, 1997.
- CHAPMAN, T.G. Stochastic modelling of daily rainfall: the impact of adjoining wet days on the distribution of rainfall amounts. *Environmental Modelling and Software* 13: 317-324, 1998.
- CHEN, J.; BRISSETTE, F. Stochastic generation of daily precipitation amounts: review and evaluation of different models. *Clim. Res.* 59, 189–206. doi:10.3354/cr01214, 2014.
- CHOWDHURY, A. F. M. K.; LOCKART, N.; WILLGOOSE, G.; KUCZERA, G.; KIEM, A. S.; PARANA MANAGE, N. Development and evaluation of a stochastic daily rainfall model with long-term variability, *Hydrol. Earth Syst. Sci.*, 21, 6541-6558, doi: 10.5194/hess-21-6541-2017, 2017.
- CONNOLLY, R. D.; SCHIRMER, J.; DUNN, P. A Daily Rainfall Disaggregation Model. *Agricultural and Forest Meteorology - AGR FOREST METEOROL.* 92. 105-117. 10.1016/S0168-1923(98)00088-4, 1998.
- COSTA, V. Estimacão indireta de quantis de enchentes extremas a partir de modelos chuva-vazão com emprego conjunto de um gerador estocástico de precipitação diária, análise bayesiana e distribuições limitadas superiormente. Tese de doutorado, Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos da UFMG, Belo Horizonte, 2015.
- COSTA, V.; FERNANDES, W.; NAGHETTINI, M. A Bayesian model for stochastic generation of daily precipitation using an upper-bounded distribution function. *Stoch Environ Res Risk Asses*, v. 29(2), pp. 563-576, 2015.



- COSTA, V.; FERNANDES, W. Bayesian estimation of extreme flood quantiles using a rainfall-runoff model and a stochastic daily rainfall generator. *J. Hydrol.*, v. 554, pp. 137-154, 2017.
- COWPERTWAIT, P. S. P. Further developments of the Neyman- Scott clustered point process for modelling rainfall. *Water Resour Res* 27(7):1431–1438, 1991.
- DE MICHELE, C.; SALVADORI, G. A generalized Pareto intensity duration model of storm rainfall exploiting 2-copulas. *J. Geophys. Res.*, 108, D24067, 2003.
- DE MICHELE, C.; SALVADORI, G.; CANOSSO, M.; PETACCIA, A.; ROSSO, R. Bivariate statistical approach to check adequacy of dam spillway, *J. Hydrol. Eng.*, 10, 50–57, 2005.
- DEIDDA, R.; BADAS, M. G.; PIGA, E. Space-time multifractality of remotely sensed rainfall fields. *J. Hydrol.*, 322, 2–13, doi:10.1016/j.jhydrol.2005.02.036, 2006.
- DIEZ-SIERRA, J.; del JESUS, M. Subdaily rainfall estimation through daily rainfall downscaling using random forests in Spain. *Water*, v. 11, W11010125, 2019.
- DINIZ, E. C. Modelos de distribuição espacial de precipitações intensas. Dissertação de mestrado, Programa de Pós-Graduação em Física, Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista, 76p, 2003.
- ECONOPOULY, T. W.; DAVIS D. R.; WOOLHISER D. A. Parameter transferability for a daily rainfall disaggregation model, *J. Hydrol.*, 118, 209–228, doi:10.1016/0022-1694(90)90259-Z, 1990.
- EVIN, G.; FAVRE, A. C.; HINGRAY, B. Stochastic generation of multi-site daily precipitation focusing on extreme events. *Hydrology and Earth System Sciences*, v. 22, n. 1, p. 655–672, 2018.
- FASANO, G.; FRANCESCHINI, A. Monthly Notices of the Royal Astronomical Society, edited, pp. 155-170, 1987.
- FERNANDES, W. S. Método para a estimação de quantis de enchentes extremas com o emprego conjunto de análise Bayesiana, de informações não sistemáticas e de distribuições limitadas superiormente. Tese de doutorado, Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos da UFMG, Belo Horizonte, 202p, 2009.
- FERNANDES, W. S.; NAGHETTINI, M.; LOSCHI, R. A Bayesian approach for estimating extreme flood probabilities with upper-bounded distribution functions. *Stoch Environ Res Risk Assess*, v. 24, pp. 1127-1143, 2010.
- FERREIRA, M. G. B. Análise regional da distribuição temporal de precipitações sub diárias para o estado de Minas Gerais. Tese de mestrado, Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos da UFMG, Belo Horizonte, 2015.
- FERRARIS, L.; GABELLANI, S.; REBORA, N.; PROVENZALE, A. A comparison of stochastic models for spatial rainfall downscaling, *Water Resour. Res.*, 39, 1368, doi:10.1029/2003WR002504, 2003.
- FURRER, E. M.; KATZ, R. W. Improving the simulation of extreme precipitation events by stochastic weather generators. *Water Resources Research*, v. 44, W12439, doi: 10.1029/2008WR007316, 2008.
- GABRIEL, K.R.; NEUMANN, J. A Markov chain model for daily rainfall occurrence at Tel Aviv. *Q. J. R. Meteorol. Soc.* 88, 90–95. doi:10.1002/qj.49708837511, 1962.
- GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. Markov Chain Monte Carlo in Practice. London: Chapman & Hall/CRC, UK, 286 p., 1996.

- GLASBEY, C. A.; COOPER, G.; MCGECHAN, M. B. Disaggregation of daily rainfall by conditional simulation from a point-process model. *Journal of Hydrology*, vol. 165, no. 1–4, pp. 1–9, 1995.
- GOYAL, M. K.; BURN, D. H.; OJHA, C. S. P. Precipitation Simulation Based on k-Nearest Neighbor Approach Using Gamma Kernel. *Journal of Hydrologic Engineering*, v. 18, n. 5, p. 481–487, 2013.
- HARROLD, T. I.; SHARMA, A.; SHEATHER, S. J. A nonparametric model for stochastic generation of daily rainfall amounts. *Water Resour. Res.*, 39(12), 1343, 2003.
- HERSHENHORN, J.; WOOLHISER D. A. Disaggregation of daily rainfall, *J. Hydrol.*, 95, 299–322, doi:10.1016/0022-1694(87)90008-4, 1987.
- HERSHFIELD, D. M. Estimating the probable maximum precipitation. Proc. ASCE, *J. Hydraul. Div.* 87(HY5), 99–106, 1961.
- HERSHFIELD, D. M. Method for estimating probable maximum precipitation. *J. Am. Waterworks Assoc.* 57, 965–972, 1965.
- HUNDECHA, Y.; PAHLOW, M.; SCHUMANN, A. Modeling of daily precipitation at multiple locations using a mixture of distributions to characterize the extremes. *Water Resources Research*, v. 45, W12412, doi: 10.1029/2008WR007453, 2009.
- KACZMARSKA, J.; ISHAM, V.; ONOF, C. Point process models for fine-resolution rainfall. *Hydrol. Sci. J.* 140806131733005. doi:10.1080/02626667.2014.925558, 2014.
- KACZMARSKA, J.; ISHAM, V.; NORTHROP, P. Local generalised method of moments: an application to point process-based rainfall models. *Environmetrics*, 26(4):312–325, 2015.
- KANG, B.; RAMÍREZ, J. A. A coupled stochastic space-time intermittent random cascade model for rainfall downscaling, *Water Resour. Res.*, 46, W10534, doi:10.1029/2008WR007692, 2010.
- KIM, D.; KWON, H. H.; LEE, S. O.; KIM, S. Regionalization of the Modified Bartlett–Lewis rectangular pulse stochastic rainfall model across the Korean Peninsula. *Journal of Hydro-environment Research*, ISSN 1570-6443, doi.org/10.1016/j.jher.2014.10.004, 2014.
- KIM, D.; CHO, H.; ONOF, C.; CHOI, M. Let-It-Rain: a web application for stochastic point rainfall generation at ungaged basins and its applicability in runoff and flood modelling. *Stochastic Environmental Research and Risk Assessment*, doi: 10.1007/s00477-016-1234-6, 2017.
- KOSSIERIS, P.; MAKROPOULOS, C.; ONOF, C.; KOUTSOYIANNIS, D. A rainfall disaggregation scheme for sub-hourly time scales: Coupling a Bartlett-Lewis based model with adjusting procedures, *Journal of Hydrology*, doi:10.1016/j.jhydrol.2016.07.015, 2016.
- KOUTSOYIANNIS, D. A. A probabilistic view of Hershfield’s method for estimating probable maximum precipitation. *Water Resources Research*, v. 35, n. 4, pp. 1313-1322, 1999.
- KOUTSOYIANNIS, D.; ONOF, C. Rainfall disaggregation using adjusting procedures on a Poisson cluster model, *Journal of Hydrology*, 246, 109–122, 2001.
- LALL, U.; SHARMA, A. A nearest neighbor bootstrap for resampling hydrological time series, *Water Resources Research*, 32, 679-693, 1996.
- LI, C.; SINGH, V. P.; MISHRA, A. K. Simulation of the entire range of daily precipitation using a hybrid probability distribution. *Water Resources Research*, v. 48, n. 3, p. 1–17, 2012.

- LI, X.; MESHGI, A.; WANG, X.; ZHANG, J.; TAY, S. H. X.; PIJCKE, G.; MANOCHA, N., ONG, M.; NGUYENB, M. T.; BABOVICA, V. Three resampling approaches based on method of fragments for daily-to-subdaily precipitation disaggregation. *Int. J. Climatol.*, 38, e1119–e1138, doi: 10.1002/joc.54, 2018.
- LICZNAR, P.; LOMOTOWSKI, J.; RUPP, D. E. Random cascade driven rainfall disaggregation for urban hydrology: an evaluation of six models and a new generator. *Atmos. Res.* 99, 563–578, 2011.
- LU, Y.; QIN, X. Multisite rainfall downscaling and disaggregation in a tropical urban area. *J. Hydrol.* 509: 55–65, doi: 10.1016/j.jhydrol.2013.11.027, 2014.
- LUNN, D.; SPIEGELHALTER, D.; THOMAS, A.; BEST, N. The BUGS project: Evolution, critique and future directions (with discussion), *Statistics in Medicine* 28: 3049—3082, 2009.
- MARAUN, D.; WETTERHALL, F.; IRESON, A. M.; CHANDLER, R. E.; KENDON, E. J.; WIDMANN, M.; BRIENEN, S.; RUST, H. W.; SAUTER, T.; THEMEßL, M.; VENEMA, V. K. C.; CHUN, K. P.; GOODESS, C. M.; JONES, R. G.; ONOF, C.; VRAC, M.; THIELE-EICH, I. Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, v. 48, i. 3. doi: 0.1029/2009rg000314, 2010.
- MATTOS, N. M. C; SILVA, R. R. Inferência bayesiana via simulação estocástica com implementação no programa WINBUGS. Simpósio Brasileiro de Pesquisa Operacional, Rio de Janeiro, 2002.
- MCINTYRE, N.; BÁRDOSSY A. Using probable maximum precipitation to bound the disaggregation of rainfall. *Water* 2017, v. 9(7), 496; doi:10.3390/w9070496, 2017.
- MEHROTRA, R.; SHARMA, A. Preserving low-frequency variability in generated daily rainfall sequences, *Journal of Hydrology*, 345, 102-120, 2007.
- MEHROTRA, R.; WESTRA, S.; SHARMA, A.; SRIKANTHAN, R. Continuous rainfall simulation: 2. A regionalized daily rainfall generation approach. *Water Resour. Res.* 48, W01536, doi:10.1029/2011WR010490, 2012.
- MEHROTRA, R.; LI, J.; WESTRA, S.; SHARMA, S. A programming tool to generate multi-site daily rainfall using a two-stage semi parametric model. *Environmental Modelling & Software*, Volume 63, Pages 230-239, ISSN 1364-8152, doi: 10.1016/j.envsoft.2014.10.016, 2015.
- MICOVIC, Z.; SCHAEFER, M. G.; TAYLOR, G. H. Uncertainty analysis for Probable Maximum Precipitation estimates. *Journal of Hydrology*, 521, 360-373p, ISSN 0022-1694, doi:10.1016/j.jhydrol.2014.12.033, 2015.
- MÜLLER, H.; HABERLANDT, U. Temporal rainfall disaggregation using a multiplicative cascade model for spatial application in urban hydrology. *Journal of Hydrology*, v. 556, 847-864. doi: 10.1016/j.jhydrol.2016.01.031, 2018.
- NAGHETTINI, M.; PINTO, E. J. A. *Hidrologia Estatística. 1a ed. Belo Horizonte: CPRM – Companhia de Pesquisa de Recursos Minerais – Superintendência Regional de Belo Horizonte*, 552 p, 2007.
- NAVEAU, P.; HUSER, R.; RIBEREAU, P.; HANNART, A., Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection, *Water Resour. Res.*, 52, 2753– 2769, 2016.

- OLSSON, J.; NIEMCZYNOWICZ, J.; BERNDTSSON, R. (1993). Fractal Analysis of High-Resolution Rainfall Time Series. *Journal of Geophysical Research*, Vol. 98, n° D12, Páginas 23,265-23,274, 20 de Dezembro de 1993.
- OLSSON, J. Evaluation of a cascade model for temporal rainfall disaggregation. *Hydrol. Earth Syst. Sci.* 2. 10.5194/hess-2-19-1998.
- ONOF, C.; CHANDLER, R. E.; A. KAKOU, NORTHROP, P. Rainfall modelling using Poisson-cluster process, paper presented at International Conference in Honour of Jacques Bernire, UNESCO, Paris, 11–13 Sept., 1995.
- ONOF, C.; CHANDLER, R. E.; KAKOU, A.; NORTHROP, P.; WHEATER, H.S.; ISHAM, V.S. Rainfall modelling using Poisson-cluster processes: a review of developments. *Stochastic Environmental Research and Risk Assessment*, 14:384–411, 2000.
- PAPALEXIOU, S. M.; KOUTOSOYIANNIS, D. A probabilistic approach to the concept of probable maximum precipitation. *Advances in Geosciences* 7, pp. 51-54, 2006.
- PAPALEXIOU, S. M.; KOUTOSOYIANNIS, D. Entropy based derivation of probability distributions: A case study to daily rainfall. *Advances in Water Resources*, v. 45, pp. 51-57, 2012.
- PAPALEXIOU, S. M.; KOUTOSOYIANNIS, D.; MACROPOULOS, C. How extreme is extreme? An assessment of daily rainfall distribution tails. *Hydro Earth Syst Scy*, v. 17, pp. 851-862, 2013.
- PAPALEXIOU, S.M.; KOUTSOYIANNIS, D. A global survey on the seasonal variation of the marginal distribution of daily precipitation. *Adv. Water Resour.* vol. 94, pp. 131–145. doi: 10.1016/j.advwatres.2016.05.005, 2016.
- PINHEIRO, M. C. Diretrizes para a Elaboração de Estudos Hidrológicos e Dimensionamentos Hidráulicos em Obras de Mineração. 1a ed. Editora: Associação Brasileira de Recursos Hídricos, ISBN: 978-85-8868-632-8, 2011.
- PODUJE, A.C.C.; HABERLANDT, U. Short time step continuous rainfall modeling and simulation of extreme events. *J. Hydrol.*, 552, pp. 182-197, 2017.
- PODUJE, A. C.C.; HABERLANDT U. Spatio-Temporal Synthesis of Continuous Precipitation Series Using Vine Copulas. *Water*. 2018; 10(7):862, 2018.
- PRESS, W. H.; TEUKOLSKY, S. A. Kolmogorov-Smirnov Test for Two-Dimensional Data *Computers in Physics* 2:4, 74-77, 1988.
- PUI, A.; SHARMA, A.; MEHROTRA, R.; SIVAKUMAR, B.; JEREMIAH, E. A comparison of alternatives for daily to sub-daily rainfall disaggregation. *J. Hydrol.* 470: 138–157, doi: 10.1016/j.jhydrol.2012.08.041, 2012.
- RAJAGOPALAN, B.; LALL, U. A k-nearest-neighbor simulator for daily precipitation and other weather variables. *Water Resour. Res.*, 35(10), 3089–3101, 1999.
- RAMESH, N. I.; GARTHWAITE, A. P.; ONOF, C. A doubly stochastic rainfall model with exponentially decaying pulses. *Stochastic Environmental Research and Risk Assessment* 32:1645–1664, 45–1664, doi: 10.1007/s00477-017-1483-z, 2018.
- ROBERT, C. P.; CASELLA, G. Monte Carlo statistical methods. New York: Springer, 2º ed., 645 p., 2004.

- RODRIGUEZ-ITURBE, I.; COX, D. F. R. S.; ISHAM, V. Some models for rainfall based on stochastic point processes, *P. R. Soc. London*, A410, 269–288, 1987.
- RUPP, D. E.; LICZNAR, P.; ADAMOWSKI, W.; LESNIEWSKI, M. Multiplicative cascade models for fine spatial downscaling of rainfall: parameterization with rain gauge data. *Hydrol. Earth Syst. Sci.*, 16, 671–684, 2012. doi:10.5194/hess-16-671-2012, 2012.
- SALAS, J. D.; LEE, T. S. Nonparametric simulation of single-site seasonal streamflows. *J. Hydrol. Eng.*, 15(4), 284–296, 2010.
- SHARIF, M.; BURN, D. H.; WEY, K. Daily and hourly weather data generation using a k-nearest neighbour approach. Proc., 18<sup>th</sup> Canadian Hydrotechnical Conf., Canadian Society for Civil Engineering, Montreal, QC, Canada, 2007.
- SHARMA, A.; LALL, U. A nonparametric approach to daily rainfall simulation, *Mathematics and Computers in Simulation*, 48, pp. 367-371, 1999.
- SHARMA, A.; SRIKANTHAN, R. Continuous rainfall simulation: a nonparametric alternative. In: 30th Hydrology and Water Resources Symposium, Launceston, Tasmania, 4–7 December, 2006.
- SHARMA, A.; MEHROTRA R. Rainfall generation, in *Rainfall: State of the Science*, edited by M. Gebremichael, pp. 215–246, American Geophysical Union, Washington, DC, 2010.
- SIVAKUMAR, B.; SHARMA, A. A cascade approach to continuous rainfall data generation at point locations. *Stochastic Environmental Research and Risk Analysis*, v. 22, p. 451-459, 2008.
- SIVAKUMAR, B. Section 3.6 - Nonparametric Methods. In: *Chaos in Hydrology: Bridging Determinism and Stochasticity*. doi; 10.1007/978-90-481-2552-4, 2017.
- SKLAR, A. Fonctions de répartition à n dimensions et leurs marges., Publications de l'Institut de Statistique de l'Université de Paris, 8, 229–231, 1959.
- SLADE, J. J. An asymmetric probability function. *Transaction: American Society of Civil Engineers*, n. 62, pp. 35-104, 1936.
- SRIKANTHAN, R.; MCMAHON, T.A. Stochastic generation of rainfall and evaporation data. AWRC Technical Paper No. 84, 301pp, 1985.
- SRIKANTHAN, R.; MCMAHON, T.A. Stochastic generation of annual, monthly and daily climate data: A review, *Hydrology and Earth System Sciences*, 5(4), pp. 653-670, 2001.
- SRIKANTHAN, R.; SIRIWARDENA, L.; MCMAHON, T.A. Comparison of two daily rainfall data generation models, *Australasian Journal of Water Resources*, 8:2, 203-212, doi: 10.1080/13241583.2005.11465257, 2005.
- SRIVASTAV, R.K.; SIMONOVIC, S.P. Multi-site, multivariate weather generator using maximum entropy bootstrap, *Clim Dyn*, 44: 3431, doi: 10.1007/s00382-014-2157-x, 2015.
- SVANIDZE, G. G. Mathematical Modeling of Hydrologic Systems, *Water Resour. Publ.*, Fort Collins, Colo, 1980.
- TAKARA, K.; LOEBIS, J. Frequency analysis introducing probable maximum hydrologic events: Preliminary studies in Japan and in Indonesia. In: LOEBIS, J. (Ed.), *International Symposium on Comparative Research on Hydrology and Water Resources in Southeast Asia and the Pacific*, 1996, Indonesian National Committee for International Hydrological Programme, pp. 67-76, 1996.

- VENEZIANO, D.; LANGOUSIS, A. Scaling and Fractals in Hydrology. In: Sivakumar, B., Berndtsson, R., (Eds.). *Advances in Data-based Approaches for Hydrologic Modeling and Forecasting*. World Scientific Publishing, Hackensack, 145p, 2010.
- VENEZIANO, D.; LANGOUSIS, A.; FURLOCO, P. Multifractality and rainfall extremes: a review. *Water Resour. Res.*, 42 (6), W06D15, 2006.
- VERDIN, A.; RAJAGOPALAN, B.; KLEIBER, W.; PODESTÁ, G., BERT, F. A conditional stochastic weather generator for seasonal to multi-decadal simulations. *J. Hydrol.*, doi.org/10.1016/j.jhydrol.2015.12.036, 2015.
- VERNIEUWE, H.; VANDENBERGHE, S.; DE BAETS, B.; VERHOEST, N. E. C. A continuous rainfall model based on vine copulas, *Hydrol. Earth Syst. Sci.*, 19, 2685-2699, <https://doi.org/10.5194/hess-19-2685-2015>, 2015.
- VICK, S. G. Degrees of belief – subjective probability and engineering judgment. *Reston: ASCE Press, USA*, 455 pp., 2002.
- VRAC, M.; NAVEAU, P. Stochastic downscaling of precipitation: From dry events to heavy rainfalls, *Water Resour. Res.*, 43, W07402, doi:10.1029/2006WR005308, 2007.
- WESTRA, S.; MEHROTRA, R.; SHARMA, A.; SRIKANTHAN, R. Continuous rainfall simulation: 1. A regionalized subdaily disaggregation approach. *Water Resour. Res.* 48, W01535, 2012.
- WHEATER, H.; CHANDLER, R.; ONOF, C.; ISHAM, V.; BELLONE, E.; YANG, C.; LEKKAS, D.; LOURMAS, G.; SEGOND, M.-L. Spatial temporal rainfall modelling for flood risk estimation, *Stochastic Environmental Research and Risk Assessment*, 19, 403–416, 2005.
- WILBY, R.L.; WIGLEY, T.M.L.; CONWAY, D.; JONES, P.D.; HEWITSON, B.C.; MAIN, J.; WILKS, D.S. Statistical downscaling of general circulation model output: A comparison of methods. *Water Resour. Res.* 34, 2995–3008. doi:10.1029/98WR02577, 1998.
- WILKS, D. S. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agric. For. Meteorol.* 93, 153–169. doi:10.1016/S0168-1923(98)00125-7, 1999.
- WILKS, D. S.; WILBY, R. L. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, 23(3), 329-357, 1999.
- WILSON, P. S.; TOUMI, R. A fundamental probability distribution for heavy rainfall, *Geophys. Res. Lett.*, 32, L14812, doi:10.1029/2005GL022465, 2005.
- WÓJCIK, R; BUIHAND, T.A. Simulation of 6-hourly rainfall and temperature by two resampling schemes. *J. Hydrol.* 273(1): 69–80, 2003.
- WOOLHISER, D. H.; ROLDÁN, J. Stochastic Daily Precipitation Models. 2. A Comparison of distribution of amounts. *Water Resources Research*, v. 18, n. 5, pp. 1461-1468, 1982.
- WORLD METEOROLOGICAL ORGANIZATION (WMO). Manual on Estimation of Probable Maximum Precipitation (PMP). 291 pp., WMO-No. 1045, ISBN 978-92-63-11045-9, 2009.

## APÊNDICE I

Tabela I.1 – Coeficientes do modelo de regressão para duração de 180 minutos

Máxima Intensidade					
Estação do ano	Intercepto ( $\beta_0$ )	Latitude ( $\beta_1$ )	Longitude ( $\beta_2$ )	LatxLong ( $\beta_3$ )	Elevação ( $\beta_4$ )
DJF	1,3692	-0,9454	-0,7300	0,3111	-0,0007
MAM	0,4315	-0,4245	-0,6882	0,3248	-0,0001
JJA	2,3737	-0,4146	-1,2252	0,3069	-0,0006
SON	0,9288	0,1989	-0,4499	-0,0438	0,0002

Tempo Intensidade máxima					
Estação do ano	Intercepto ( $\beta_0$ )	Latitude ( $\beta_1$ )	Longitude ( $\beta_2$ )	LatxLong ( $\beta_3$ )	Elevação ( $\beta_4$ )
DJF	-26,5661	0,0000	0,0000	0,0000	0,0000
MAM	-3,7642	-0,1715	-0,3550	0,0584	-0,0009
JJA	-0,2821	0,6195	-0,5347	-0,1660	-0,0007
SON	-3,4161	0,2574	-0,4279	-0,2286	-0,0027

Fração de zeros					
Estação do ano	Intercepto ( $\beta_0$ )	Latitude ( $\beta_1$ )	Longitude ( $\beta_2$ )	LatxLong ( $\beta_3$ )	Elevação ( $\beta_4$ )
DJF	-26,5661	0,0000	0,0000	0,0000	0,0000
MAM	-26,5661	0,0000	0,0000	0,0000	0,0000
JJA	-2,9265	0,7723	-0,0828	-0,3298	-0,0008
SON	-26,5661	0,0000	0,0000	0,0000	0,0000

**Tabela I.2 – Coeficientes do modelo de regressão para duração de 360 minutos****Máxima Intensidade**

<b>Estação do ano</b>	<b>Intercepto (<math>\beta_0</math>)</b>	<b>Latitude (<math>\beta_1</math>)</b>	<b>Longitude (<math>\beta_2</math>)</b>	<b>LatxLong (<math>\beta_3</math>)</b>	<b>Elevação (<math>\beta_4</math>)</b>
<b>DJF</b>	1,4108	-1,0828	-0,7838	0,4526	-0,0009
<b>MAM</b>	0,3760	-0,2652	-0,6430	0,3037	-0,0003
<b>JJA</b>	2,0095	-0,2586	-0,9016	0,1726	-0,0007
<b>SON</b>	1,0502	0,0903	-0,4713	-0,0203	0,0001

**Tempo Intensidade máxima**

<b>Estação do ano</b>	<b>Intercepto (<math>\beta_0</math>)</b>	<b>Latitude (<math>\beta_1</math>)</b>	<b>Longitude (<math>\beta_2</math>)</b>	<b>LatxLong (<math>\beta_3</math>)</b>	<b>Elevação (<math>\beta_4</math>)</b>
<b>DJF</b>	1,4108	-1,0828	-0,7838	0,4526	-0,0009
<b>MAM</b>	0,3760	-0,2652	-0,6430	0,3037	-0,0003
<b>JJA</b>	2,0095	-0,2586	-0,9016	0,1726	-0,0007
<b>SON</b>	1,0502	0,0903	-0,4713	-0,0203	0,0001

**Fração de zeros**

<b>Estação do ano</b>	<b>Intercepto (<math>\beta_0</math>)</b>	<b>Latitude (<math>\beta_1</math>)</b>	<b>Longitude (<math>\beta_2</math>)</b>	<b>LatxLong (<math>\beta_3</math>)</b>	<b>Elevação (<math>\beta_4</math>)</b>
<b>DJF</b>	-26,5661	0,0000	0,0000	0,0000	0,0000
<b>MAM</b>	-26,5661	0,0000	0,0000	0,0000	0,0000
<b>JJA</b>	-4,3349	1,4737	0,0179	-0,3705	-0,0008
<b>SON</b>	-26,5661	0,0000	0,0000	0,0000	0,0000



**Tabela I.3 – Coeficientes do modelo de regressão para duração de 720 minutos**

<b>Máxima Intensidade</b>					
<b>Estação do ano</b>	<b>Intercepto (<math>\beta_0</math>)</b>	<b>Latitude (<math>\beta_1</math>)</b>	<b>Longitude (<math>\beta_2</math>)</b>	<b>LatxLong (<math>\beta_3</math>)</b>	<b>Elevação (<math>\beta_4</math>)</b>
<b>DJF</b>	1,4855	-1,2396	-0,8775	0,4863	-0,0005
<b>MAM</b>	0,1685	-0,1276	-0,6044	0,2245	0,0000
<b>JJA</b>	1,6810	0,0011	-0,8067	0,0768	-0,0005
<b>SON</b>	0,9775	0,2184	-0,4137	-0,0725	0,0001

<b>Tempo Intensidade máxima</b>					
<b>Estação do ano</b>	<b>Intercepto (<math>\beta_0</math>)</b>	<b>Latitude (<math>\beta_1</math>)</b>	<b>Longitude (<math>\beta_2</math>)</b>	<b>LatxLong (<math>\beta_3</math>)</b>	<b>Elevação (<math>\beta_4</math>)</b>
<b>DJF</b>	-26,5661	0,0000	0,0000	0,0000	0,0000
<b>MAM</b>	-26,5661	0,0000	0,0000	0,0000	0,0000
<b>JJA</b>	-26,5661	0,0000	0,0000	0,0000	0,0000
<b>SON</b>	-26,5661	0,0000	0,0000	0,0000	0,0000

<b>Fração de zeros</b>					
<b>Estação do ano</b>	<b>Intercepto (<math>\beta_0</math>)</b>	<b>Latitude (<math>\beta_1</math>)</b>	<b>Longitude (<math>\beta_2</math>)</b>	<b>LatxLong (<math>\beta_3</math>)</b>	<b>Elevação (<math>\beta_4</math>)</b>
<b>DJF</b>	-26,5661	0,0000	0,0000	0,0000	0,0000
<b>MAM</b>	-26,5661	0,0000	0,0000	0,0000	0,0000
<b>JJA</b>	-26,5661	0,0000	0,0000	0,0000	0,0000
<b>SON</b>	-26,5661	0,0000	0,0000	0,0000	0,0000

**Tabela I.4 – Probabilidades de similaridade para o atributo hidrológico INTENSIDADE MÁXIMA para cada estação do ano e duração de 60 minutos**

DJF		MAM		JJA		SON	
Estação	Prob.	Estação	Prob.	Estação	Prob.	Estação	Prob.
01943009	0,630	02043002	0,344	01843002	0,788	01843002	0,650
01943022	0,624	01943022	0,340	02043013	0,768	01844009	0,596
02044024	0,616	01943035	0,337	01943002	0,750	01943002	0,594
01944009	0,614	01943009	0,323	02044007	0,746	01943035	0,593
01944004	0,613	01843002	0,323	01943035	0,745	01844001	0,590
01944027	0,611	01943002	0,319	01844009	0,738	02043013	0,587
01944062	0,604	02043013	0,317	02043010	0,721	01844010	0,573
02044052	0,597	02043010	0,311	01844001	0,718	02044007	0,570
02044054	0,596	01944009	0,308	01943009	0,714	02043010	0,569
02043013	0,591	01844009	0,306	01944009	0,712	01943009	0,544
02044021	0,581	01844010	0,299	01844010	0,708	01944009	0,537
02044041	0,572	02044007	0,297	02044024	0,682	02043002	0,530
01943035	0,543	01844001	0,296	02044054	0,681	01943022	0,525
02043002	0,541	01944004	0,289	02044052	0,673	02044024	0,502
02044007	0,534	01944027	0,285	02044021	0,670	02044021	0,499
01943002	0,499	02044024	0,282	01944062	0,669	02044054	0,498
02043010	0,496	02044021	0,278	01943022	0,666	01944004	0,497
01944049	0,490	02044052	0,275	01944027	0,666	02044052	0,496
01944021	0,479	01942032	0,274	01944004	0,666	01944027	0,496
02044042	0,452	02044054	0,271	02044041	0,661	02044041	0,491
02045013	0,441	02044041	0,270	02043002	0,640	01944062	0,487
01942031	0,419	01942031	0,270	01944049	0,636	01944049	0,478
01844009	0,401	01942030	0,270	02044042	0,631	02044042	0,470
01843002	0,394	01944062	0,266	01944021	0,574	01944021	0,441
01844010	0,384	01944049	0,254	02045013	0,534	01942032	0,440
02045002	0,363	01942008	0,243	01845021	0,510	01942030	0,433
01942030	0,337	01944021	0,243	01942031	0,489	01942031	0,428
02045012	0,334	02044042	0,242	01942032	0,484	02045013	0,407
01942008	0,327	01841011	0,232	01942030	0,482	01942008	0,404
01844001	0,318	01845021	0,226	02045002	0,463	01845021	0,402
01942032	0,317	01941005	0,223	02045012	0,461	02045002	0,366
01940020	0,285	02045013	0,213	01942008	0,457	02045012	0,349
01946009	0,274	01941006	0,204	01845004	0,415	01845004	0,349
01845004	0,254	01845004	0,203	01946009	0,386	01841011	0,347
01941006	0,250	02045002	0,190	01940020	0,368	01941005	0,335
01845021	0,230	01946009	0,182	01841011	0,336	01941006	0,331
01941012	0,221	01940009	0,181	01941006	0,328	01946009	0,328
01841011	0,206	01941012	0,178	01941012	0,327	01940020	0,297
01941005	0,206	02045012	0,176	01941005	0,322	01941012	0,296
01940009	0,161	01940020	0,133	01940009	0,227	01940009	0,265

**Tabela I.5 – Probabilidades de similaridade para o atributo hidrológico TEMPO DA INTENSIDADE MÁXIMA para cada estação do ano e duração de 60 minutos**

DJF		MAM		JJA		SON	
Estação	Prob.	Estação	Prob.	Estação	Prob.	Estação	Prob.
01845021	0,035	01845021	0,061	01843002	0,581	01843002	0,186
02044007	0,022	01943035	0,060	01844009	0,507	02043013	0,169
02043013	0,021	02043013	0,059	02043013	0,506	02044007	0,169
02044042	0,018	01843002	0,058	01844001	0,505	01844001	0,168
02044054	0,016	01943002	0,056	01943002	0,498	01844009	0,167
01843002	0,015	01844001	0,055	02044007	0,490	01943002	0,163
01944062	0,015	01943009	0,055	01943035	0,475	01844010	0,157
01844001	0,014	01844009	0,054	01844010	0,470	02043010	0,152
02045012	0,014	01943022	0,054	02043010	0,456	01943035	0,151
01844009	0,014	02043010	0,052	01944009	0,417	02044042	0,145
02044024	0,013	02044007	0,052	01943009	0,413	01944009	0,144
01943002	0,013	02043002	0,052	02044054	0,383	02044054	0,143
01944049	0,013	01844010	0,052	02044024	0,378	01845021	0,143
02044052	0,013	01944009	0,050	02044052	0,372	01944062	0,140
01944009	0,013	02044024	0,039	02044021	0,370	02044024	0,139
02044041	0,012	01944004	0,039	01944062	0,366	01943009	0,139
02044021	0,012	01944027	0,038	02044042	0,366	02044052	0,139
01844010	0,012	02044021	0,038	02044041	0,363	01944049	0,138
02043010	0,011	02044054	0,037	01944027	0,357	02044021	0,137
01943009	0,010	02044052	0,037	01944049	0,357	02044041	0,137
01944027	0,010	02044041	0,036	01944004	0,356	01944027	0,132
01943035	0,010	01944062	0,034	01943022	0,348	01944004	0,130
01944004	0,010	02044042	0,033	02043002	0,333	02045012	0,120
01845004	0,009	01944049	0,033	01845021	0,331	01944021	0,119
02045013	0,007	01942032	0,031	01944021	0,288	02045013	0,117
01944021	0,007	01942030	0,029	02045013	0,260	01943022	0,116
02045002	0,006	01845004	0,028	01942032	0,238	02043002	0,109
01941012	0,005	01841011	0,025	02045012	0,232	01845004	0,108
01940020	0,005	01944021	0,025	01942030	0,230	02045002	0,108
01946009	0,005	01942031	0,024	01845004	0,219	01940020	0,102
01943022	0,004	01942008	0,023	02045002	0,217	01946009	0,099
02043002	0,003	01941005	0,022	01942031	0,215	01942008	0,094
01942008	0,002	02045013	0,018	01942008	0,215	01942032	0,094
01942032	0,002	01941012	0,018	01946009	0,182	01942030	0,093
01942030	0,002	02045012	0,016	01940020	0,159	01941012	0,092
01942031	0,001	01946009	0,015	01841011	0,158	01942031	0,088
01841011	0,001	02045002	0,014	01941012	0,158	01841011	0,078
01941005	0,001	01940009	0,012	01941005	0,147	01941005	0,076
01941006	0,001	01941006	0,012	01941006	0,135	01941006	0,074
01940009	0,001	01940020	0,005	01940009	0,097	01940009	0,064

**Tabela I.6 – Probabilidades de similaridade para o atributo hidrológico INTENSIDADE MÁXIMA para cada estação do ano e duração de 180 minutos**

DJF		MAM		JJA		SON	
Estação	Prob.	Estação	Prob.	Estação	Prob.	Estação	Prob.
01943022	0,712	01943022	0,556	02043013	0,872	01843002	0,773
01943009	0,708	01943009	0,542	01943009	0,870	01943035	0,745
01944009	0,689	02043002	0,526	01943022	0,867	01943002	0,737
01944004	0,686	02043013	0,521	01943035	0,865	01844009	0,736
02044024	0,685	01943035	0,520	01944009	0,854	01844001	0,732
01944027	0,683	01944009	0,519	01943002	0,843	02043013	0,727
01944062	0,670	01944004	0,493	02043002	0,839	01844010	0,722
02044052	0,666	01944027	0,488	02044007	0,836	02043010	0,720
02044054	0,663	01943002	0,487	02043010	0,827	02043002	0,720
02043013	0,660	02044024	0,487	01843002	0,825	02044007	0,709
02044021	0,652	02044007	0,482	02044024	0,820	01943022	0,709
02044041	0,642	02043010	0,480	01944004	0,819	01943009	0,707
02043002	0,636	02044052	0,473	01944027	0,815	01944009	0,694
01943035	0,620	02044021	0,470	02044052	0,804	01944004	0,661
02044007	0,603	02044054	0,469	02044054	0,803	01944027	0,658
02043010	0,571	01944062	0,463	02044021	0,800	02044024	0,658
01943002	0,570	02044041	0,458	01844009	0,800	02044021	0,656
01944049	0,561	01843002	0,445	01944062	0,793	02044052	0,652
01944021	0,547	01844009	0,442	02044041	0,785	02044054	0,650
02044042	0,523	01844010	0,434	01844010	0,778	02044041	0,647
01942031	0,500	01944049	0,416	01844001	0,755	01942032	0,641
02045013	0,498	01844001	0,409	01944049	0,729	01944062	0,639
01844009	0,466	02044042	0,396	02044042	0,702	01942030	0,633
01844010	0,454	01942031	0,391	01944021	0,670	01944049	0,632
01843002	0,445	01944021	0,387	01942031	0,639	01942031	0,625
01942030	0,421	01942032	0,379	01942032	0,614	02044042	0,620
02045002	0,414	01942030	0,377	01942030	0,612	01944021	0,604
01942008	0,403	01845021	0,374	01845021	0,592	01942008	0,594
01942032	0,402	01942008	0,344	02045013	0,576	01845021	0,579
02045012	0,392	02045013	0,335	01942008	0,549	02045013	0,560
01844001	0,380	01845004	0,314	01845004	0,464	01841011	0,559
01845004	0,340	01841011	0,309	02045002	0,453	01941005	0,544
01946009	0,336	01941005	0,295	02045012	0,433	01845004	0,524
01845021	0,333	02045002	0,283	01841011	0,422	01941006	0,520
01941006	0,306	02045012	0,277	01941005	0,393	02045002	0,517
01940020	0,296	01946009	0,264	01946009	0,376	02045012	0,490
01941012	0,294	01941012	0,260	01941006	0,345	01946009	0,490
01841011	0,286	01941006	0,258	01941012	0,333	01941012	0,465
01941005	0,282	01940009	0,223	01940009	0,230	01940009	0,457
01940009	0,219	01940020	0,173	01940020	0,221	01940020	0,411

**Tabela I.7 – Probabilidades de similaridade para o atributo hidrológico TEMPO DA INTENSIDADE MÁXIMA para cada estação do ano e duração de 180 minutos**

DJF		MAM		JJA		SON	
Estação	Prob.	Estação	Prob.	Estação	Prob.	Estação	Prob.
01844001	0,000	02043013	0,019	01843002	0,589	02043013	0,030
01845021	0,000	01943009	0,018	01844001	0,515	02044007	0,027
02044042	0,000	01944009	0,018	01844009	0,511	02044054	0,023
01940009	0,000	02044024	0,017	01943002	0,496	01944062	0,022
01940020	0,000	02044007	0,017	02043013	0,490	02044024	0,022
01942008	0,000	02044054	0,017	01943035	0,479	01944009	0,022
01942030	0,000	01944062	0,017	01844010	0,476	02044052	0,021
01942032	0,000	02044052	0,017	02044007	0,472	01843002	0,020
01946009	0,000	01944027	0,016	02043010	0,454	01943002	0,020
01943035	0,000	01944004	0,016	01943009	0,403	01943009	0,020
01844010	0,000	02044021	0,016	01944009	0,400	02044021	0,019
01844009	0,000	01943035	0,016	02044054	0,354	02044041	0,019
01843002	0,000	02044041	0,016	01943022	0,353	01944027	0,018
02044007	0,000	01943022	0,015	02044024	0,352	01943035	0,018
02043013	0,000	01943002	0,015	02043002	0,352	01944004	0,018
02045012	0,000	02043010	0,015	02044021	0,348	02044042	0,018
02045002	0,000	01944049	0,014	02044052	0,346	02043010	0,017
02045013	0,000	01843002	0,014	02044042	0,339	01844009	0,017
01944021	0,000	02044042	0,014	02044041	0,338	01944049	0,016
01944049	0,000	01844009	0,013	01944004	0,336	01844010	0,014
01845004	0,000	02043002	0,013	01944027	0,336	01844001	0,014
01841011	0,000	01844010	0,012	01944062	0,334	01943022	0,012
01941005	0,000	01944021	0,012	01944049	0,334	01944021	0,011
01941006	0,000	01844001	0,012	01845021	0,329	02045013	0,011
01941012	0,000	02045013	0,011	01944021	0,269	02045012	0,010
01942031	0,000	01845021	0,010	01942032	0,255	01845021	0,010
01943002	0,000	02045012	0,010	01942030	0,244	02043002	0,008
02043010	0,000	02045002	0,009	02045013	0,234	02045002	0,008
02043002	0,000	01942031	0,009	01942031	0,222	01940020	0,006
01943022	0,000	01845004	0,008	01942008	0,219	01845004	0,006
01943009	0,000	01942030	0,008	01845004	0,212	01946009	0,005
01944009	0,000	01942008	0,008	02045012	0,202	01942031	0,004
02044052	0,000	01942032	0,008	02045002	0,193	01942008	0,004
01944027	0,000	01946009	0,008	01841011	0,173	01942030	0,004
01944004	0,000	01940020	0,007	01946009	0,168	01941012	0,004
01944062	0,000	01941012	0,007	01941005	0,160	01942032	0,003
02044021	0,000	01841011	0,006	01941012	0,150	01941006	0,002
02044024	0,000	01941006	0,006	01941006	0,137	01841011	0,002
02044041	0,000	01941005	0,006	01940020	0,127	01941005	0,002
02044054	0,000	01940009	0,004	01940009	0,102	01940009	0,001

**Tabela I.8 – Probabilidades de similaridade para o atributo hidrológico INTENSIDADE MÁXIMA para cada estação do ano e duração de 360 minutos**

DJF		MAM		JJA		SON	
Estação	Prob.	Estação	Prob.	Estação	Prob.	Estação	Prob.
01943022	0,709	01943009	0,533	02043013	0,844	01843002	0,763
01943009	0,706	01943022	0,532	01943009	0,837	01943035	0,750
01944009	0,686	02043013	0,531	01943035	0,832	01943002	0,739
02044024	0,684	01943035	0,523	01943022	0,828	02043013	0,737
01944004	0,684	01944009	0,516	01944009	0,825	01844009	0,730
01944027	0,681	02043002	0,505	01943002	0,815	02043002	0,729
01944062	0,669	01943002	0,502	02044007	0,813	01943022	0,726
02044052	0,664	02044007	0,499	01843002	0,803	02043010	0,724
02044054	0,662	02043010	0,491	02043010	0,800	01943009	0,724
02043013	0,650	02044024	0,483	02043002	0,798	01844001	0,719
02044021	0,649	01944004	0,483	02044024	0,797	02044007	0,716
02044041	0,638	01944027	0,479	01944004	0,794	01844010	0,716
02043002	0,621	01843002	0,477	01944027	0,792	01944009	0,711
01943035	0,600	02044052	0,470	02044054	0,785	01944004	0,682
02044007	0,592	02044054	0,470	02044052	0,784	01944027	0,679
01944049	0,556	01844009	0,470	02044021	0,780	02044024	0,679
02043010	0,554	02044021	0,469	01844009	0,779	02044021	0,674
01943002	0,548	01944062	0,461	01944062	0,778	02044052	0,672
01944021	0,537	01844010	0,459	02044041	0,769	02044054	0,670
02044042	0,522	02044041	0,457	01844010	0,759	02044041	0,665
02045013	0,486	01844001	0,447	01844001	0,742	01944062	0,661
01942031	0,477	01845021	0,433	01944049	0,725	01944049	0,644
01844009	0,438	01944049	0,426	02044042	0,706	01942032	0,638
01844010	0,430	02044042	0,416	01944021	0,676	01942031	0,634
01942030	0,402	01944021	0,387	01942031	0,635	01942030	0,633
02045002	0,402	01942032	0,377	01942032	0,611	02044042	0,630
01843002	0,402	01942030	0,372	01942030	0,611	01944021	0,619
02045012	0,397	01942031	0,372	02045013	0,611	01942008	0,596
01942008	0,389	01942008	0,343	01845021	0,606	02045013	0,576
01942032	0,384	01845004	0,342	01942008	0,569	01845021	0,569
01845021	0,372	02045013	0,339	02045002	0,517	01841011	0,551
01845004	0,358	01841011	0,315	01845004	0,509	01941005	0,537
01844001	0,353	01941005	0,300	02045012	0,505	02045002	0,529
01946009	0,336	02045012	0,298	01841011	0,455	01845004	0,524
01941012	0,307	02045002	0,290	01946009	0,447	01941006	0,521
01941006	0,284	01941012	0,280	01941005	0,434	02045012	0,501
01841011	0,284	01946009	0,279	01941006	0,406	01946009	0,495
01941005	0,279	01941006	0,252	01941012	0,403	01941012	0,466
01940020	0,272	01940009	0,227	01940020	0,332	01940009	0,451
01940009	0,215	01940020	0,182	01940009	0,299	01940020	0,422

**Tabela I.9 – Probabilidades de similaridade para o atributo hidrológico TEMPO DA INTENSIDADE MÁXIMA para cada estação do ano e duração de 360 minutos**

DJF		MAM		JJA		SON	
Estação	Prob.	Estação	Prob.	Estação	Prob.	Estação	Prob.
01844001	0,000	01843002	0,008	01843002	0,396	01844001	0,000
01845021	0,000	02043013	0,007	01844001	0,348	01845021	0,000
02044042	0,000	01943035	0,005	01844009	0,336	02044042	0,000
01940009	0,000	01943002	0,004	01943002	0,323	01940009	0,000
01940020	0,000	02044007	0,004	01943035	0,320	01940020	0,000
01942008	0,000	01943009	0,003	02043013	0,314	01942008	0,000
01942030	0,000	01844009	0,003	01844010	0,312	01942030	0,000
01942032	0,000	01944009	0,003	02044007	0,293	01942032	0,000
01946009	0,000	02043010	0,003	02043010	0,291	01946009	0,000
01943035	0,000	01844010	0,002	01943009	0,257	01943035	0,000
01844010	0,000	01844001	0,002	01845021	0,254	01844010	0,000
01844009	0,000	01943022	0,002	01944009	0,245	01844009	0,000
01843002	0,000	02044024	0,002	02043002	0,238	01843002	0,000
02044007	0,000	02044054	0,001	01943022	0,233	02044007	0,000
02043013	0,000	02044052	0,001	02044024	0,196	02043013	0,000
02045012	0,000	01944004	0,001	02044021	0,194	02045012	0,000
02045002	0,000	01944027	0,001	02044054	0,193	02045002	0,000
02045013	0,000	02044021	0,001	01944004	0,191	02045013	0,000
01944021	0,000	01944062	0,001	02044052	0,190	01944021	0,000
01944049	0,000	02044041	0,001	01944027	0,188	01944049	0,000
01845004	0,000	02043002	0,001	02044042	0,184	01845004	0,000
01841011	0,000	01944049	0,001	02044041	0,184	01841011	0,000
01941005	0,000	02044042	0,000	01944049	0,181	01941005	0,000
01941006	0,000	01944021	0,000	01944062	0,176	01941006	0,000
01941012	0,000	02045013	0,000	01942032	0,158	01941012	0,000
01942031	0,000	01942031	0,000	01942030	0,145	01942031	0,000
01943002	0,000	01942030	0,000	01944021	0,133	01943002	0,000
02043010	0,000	01942032	0,000	01845004	0,128	02043010	0,000
02043002	0,000	01845021	0,000	01942031	0,120	02043002	0,000
01943022	0,000	01942008	0,000	01942008	0,120	01943022	0,000
01943009	0,000	02045002	0,000	01841011	0,111	01943009	0,000
01944009	0,000	02045012	0,000	02045013	0,103	01944009	0,000
02044052	0,000	01845004	0,000	01941005	0,098	02044052	0,000
01944027	0,000	01946009	0,000	02045012	0,088	01944027	0,000
01944004	0,000	01940020	0,000	01941012	0,080	01944004	0,000
01944062	0,000	01941006	0,000	02045002	0,080	01944062	0,000
02044021	0,000	01841011	0,000	01946009	0,079	02044021	0,000
02044024	0,000	01941005	0,000	01941006	0,064	02044024	0,000
02044041	0,000	01941012	0,000	01940009	0,054	02044041	0,000
02044054	0,000	01940009	0,000	01940020	0,036	02044054	0,000

**Tabela I. 10 – Média das probabilidades de similaridade encontradas por estação do ano  
– Duração de 180 min**

DJF		MAM		JJA		SON	
Estação	Prob.	Estação	Prob.	Estação	Prob.	Estação	Prob.
01943022	0,356	01943022	0,285	01843002	0,707	01843002	0,397
01943009	0,354	01943009	0,280	02043013	0,681	01943035	0,381
01944009	0,344	02043013	0,270	01943035	0,672	01943002	0,379
01944004	0,343	02043002	0,269	01943002	0,669	02043013	0,378
02044024	0,343	01944009	0,268	01844009	0,655	01844009	0,377
01944027	0,342	01943035	0,268	02044007	0,654	01844001	0,373
01944062	0,335	01944004	0,255	02043010	0,640	02043010	0,369
02044052	0,333	01944027	0,252	01943009	0,637	01844010	0,368
02044054	0,332	02044024	0,252	01844001	0,635	02044007	0,368
02043013	0,330	01943002	0,251	01944009	0,627	02043002	0,364
02044021	0,326	02044007	0,249	01844010	0,627	01943009	0,363
02044041	0,321	02043010	0,247	01943022	0,610	01943022	0,360
02043002	0,318	02044052	0,245	02043002	0,595	01944009	0,358
01943035	0,310	02044021	0,243	02044024	0,586	02044024	0,340
02044007	0,302	02044054	0,243	02044054	0,578	01944004	0,340
02043010	0,286	01944062	0,240	01944004	0,578	01944027	0,338
01943002	0,285	02044041	0,237	01944027	0,576	02044021	0,338
01944049	0,280	01843002	0,229	02044052	0,575	02044054	0,336
01944021	0,273	01844009	0,228	02044021	0,574	02044052	0,336
02044042	0,262	01844010	0,223	01944062	0,564	02044041	0,333
01942031	0,250	01944049	0,215	02044041	0,562	01944062	0,331
02045013	0,249	01844001	0,210	01944049	0,532	01944049	0,324
01844009	0,233	02044042	0,205	02044042	0,520	01942032	0,322
01844010	0,227	01942031	0,200	01944021	0,469	02044042	0,319
01843002	0,223	01944021	0,200	01845021	0,460	01942030	0,318
01942030	0,210	01942032	0,193	01942032	0,434	01942031	0,315
02045002	0,207	01942030	0,193	01942031	0,430	01944021	0,308
01942008	0,202	01845021	0,192	01942030	0,428	01942008	0,299
01942032	0,201	01942008	0,176	02045013	0,405	01845021	0,294
02045012	0,196	02045013	0,173	01942008	0,384	02045013	0,286
01844001	0,190	01845004	0,161	01845004	0,338	01841011	0,281
01845004	0,170	01841011	0,157	02045002	0,323	01941005	0,273
01946009	0,168	01941005	0,150	02045012	0,318	01845004	0,265
01845021	0,167	02045002	0,146	01841011	0,297	02045002	0,262
01941006	0,153	02045012	0,143	01941005	0,276	01941006	0,261
01940020	0,148	01946009	0,136	01946009	0,272	02045012	0,250
01941012	0,147	01941012	0,134	01941012	0,241	01946009	0,247
01841011	0,143	01941006	0,132	01941006	0,241	01941012	0,235
01941005	0,141	01940009	0,114	01940020	0,174	01940009	0,229
01940009	0,110	01940020	0,090	01940009	0,166	01940020	0,209

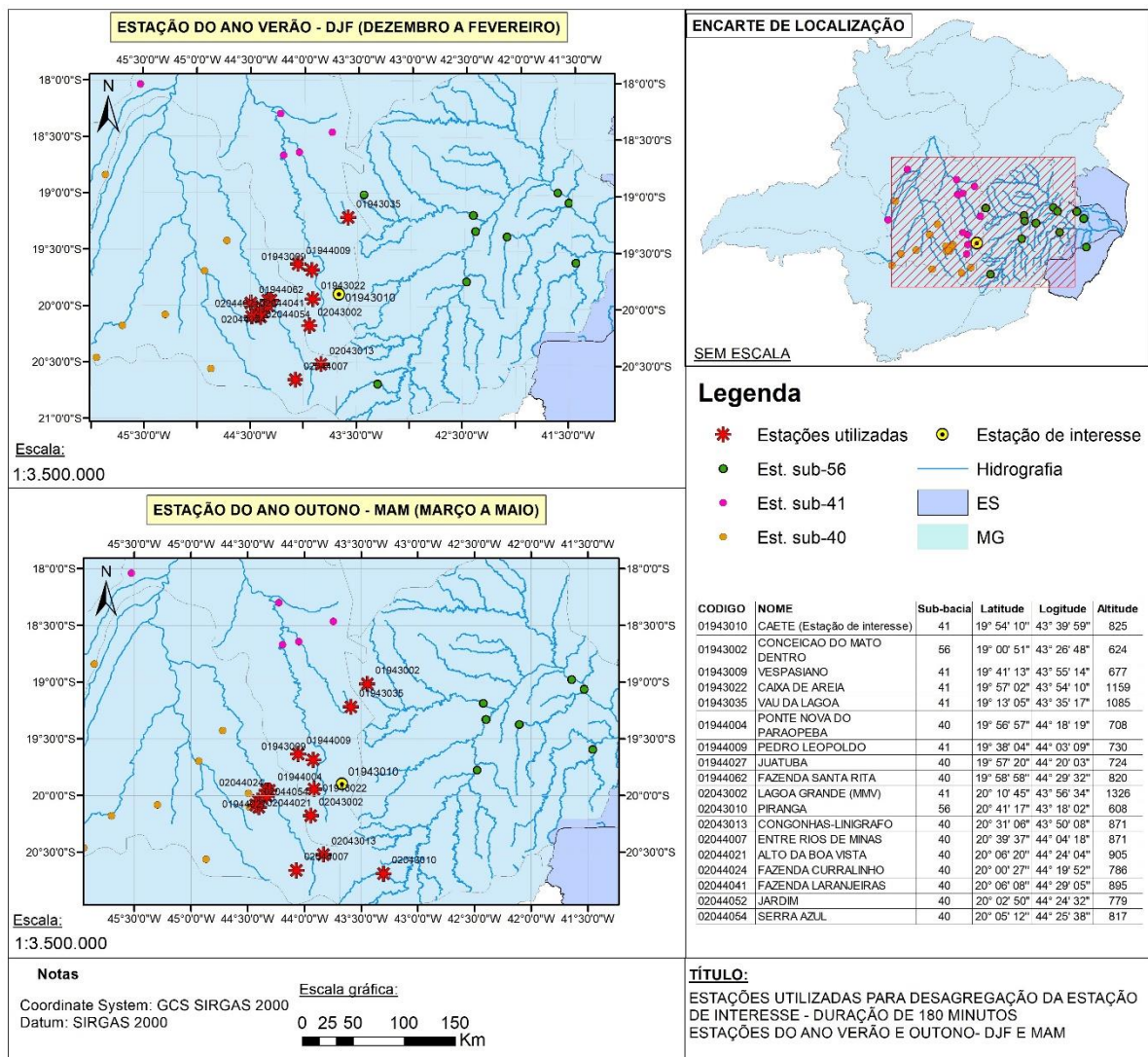


**Tabela I.11 – Média das probabilidades de similaridade encontradas por estação do ano  
– Duração de 360 min**

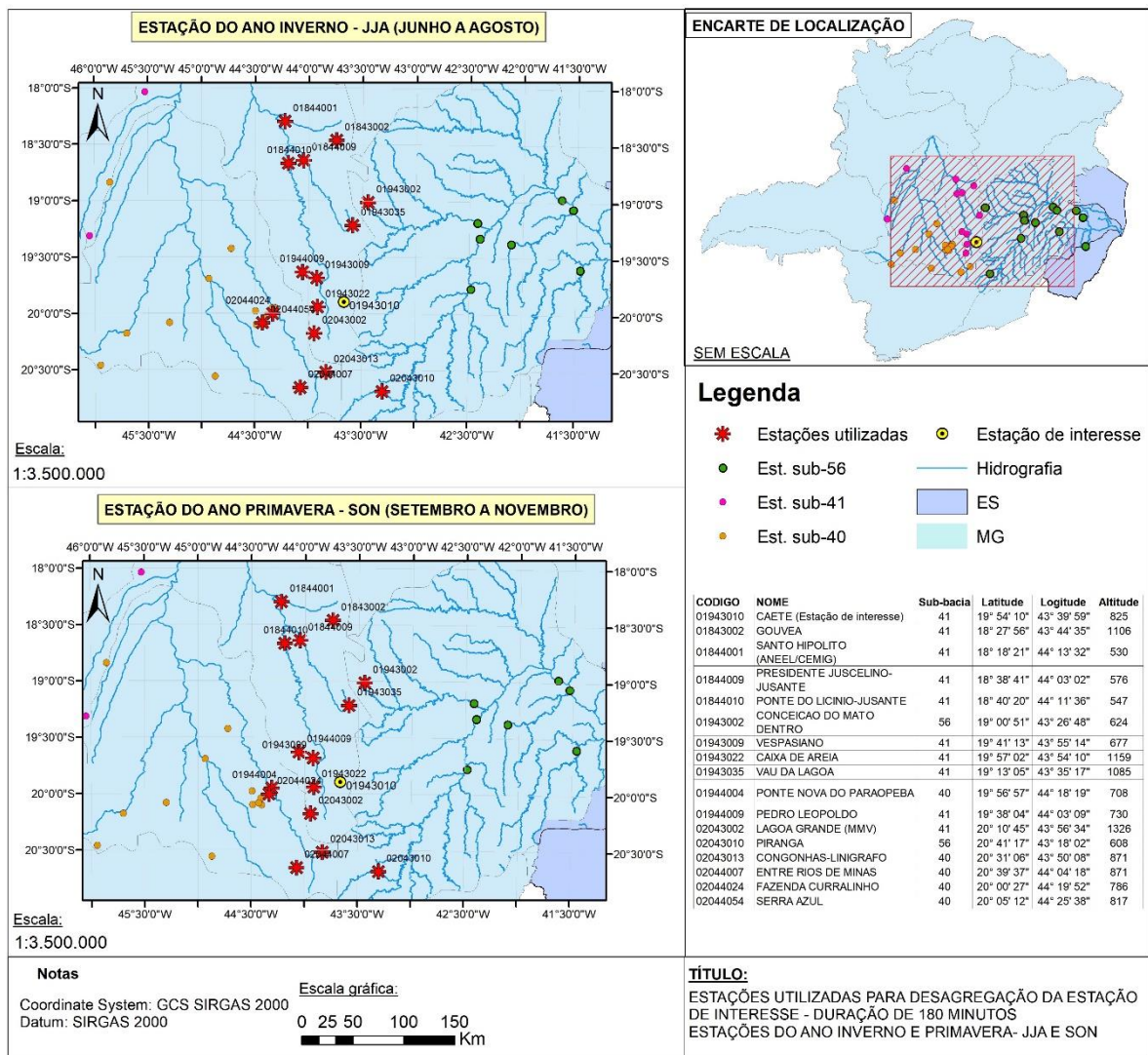
DJF		MAM		JJA		SON	
Estação	Prob.	Estação	Prob.	Estação	Prob.	Estação	Prob.
01943022	0,354	02043013	0,269	01843002	0,600	01843002	0,381
01943009	0,353	01943009	0,268	02043013	0,579	01943035	0,375
01944009	0,343	01943022	0,267	01943035	0,576	01943002	0,370
02044024	0,342	01943035	0,264	01943002	0,569	02043013	0,368
01944004	0,342	01944009	0,259	01844009	0,557	01844009	0,365
01944027	0,341	01943002	0,253	02044007	0,553	02043002	0,365
01944062	0,334	02043002	0,253	01943009	0,547	01943022	0,363
02044052	0,332	02044007	0,251	02043010	0,545	02043010	0,362
02044054	0,331	02043010	0,247	01844001	0,545	01943009	0,362
02043013	0,325	02044024	0,242	01844010	0,535	01844001	0,360
02044021	0,324	01843002	0,242	01944009	0,535	02044007	0,358
02044041	0,319	01944004	0,242	01943022	0,531	01844010	0,358
02043002	0,311	01944027	0,240	02043002	0,518	01944009	0,355
01943035	0,300	01844009	0,236	02044024	0,497	01944004	0,341
02044007	0,296	02044052	0,236	01944004	0,492	01944027	0,339
01944049	0,278	02044054	0,236	01944027	0,490	02044024	0,339
02043010	0,277	02044021	0,235	02044054	0,489	02044021	0,337
01943002	0,274	01944062	0,231	02044021	0,487	02044052	0,336
01944021	0,268	01844010	0,230	02044052	0,487	02044054	0,335
02044042	0,261	02044041	0,229	01944062	0,477	02044041	0,333
02045013	0,243	01844001	0,224	02044041	0,476	01944062	0,330
01942031	0,239	01845021	0,217	01944049	0,453	01944049	0,322
01844009	0,219	01944049	0,213	02044042	0,445	01942032	0,319
01844010	0,215	02044042	0,208	01845021	0,430	01942031	0,317
01942030	0,201	01944021	0,193	01944021	0,405	01942030	0,316
02045002	0,201	01942032	0,189	01942032	0,385	02044042	0,315
01843002	0,201	01942030	0,186	01942030	0,378	01944021	0,309
02045012	0,199	01942031	0,186	01942031	0,378	01942008	0,298
01942008	0,194	01942008	0,172	02045013	0,357	02045013	0,288
01942032	0,192	01845004	0,171	01942008	0,344	01845021	0,284
01845021	0,186	02045013	0,170	01845004	0,318	01841011	0,276
01845004	0,179	01841011	0,158	02045002	0,298	01941005	0,268
01844001	0,176	01941005	0,150	02045012	0,297	02045002	0,264
01946009	0,168	02045012	0,149	01841011	0,283	01845004	0,262
01941012	0,153	02045002	0,145	01941005	0,266	01941006	0,261
01941006	0,142	01941012	0,140	01946009	0,263	02045012	0,250
01841011	0,142	01946009	0,139	01941012	0,242	01946009	0,247
01941005	0,139	01941006	0,126	01941006	0,235	01941012	0,233
01940020	0,136	01940009	0,113	01940020	0,184	01940009	0,226
01940009	0,108	01940020	0,091	01940009	0,177	01940020	0,211

**Tabela I.12 – Média das probabilidades de similaridade encontradas por estação do ano  
– Duração de 720 min**

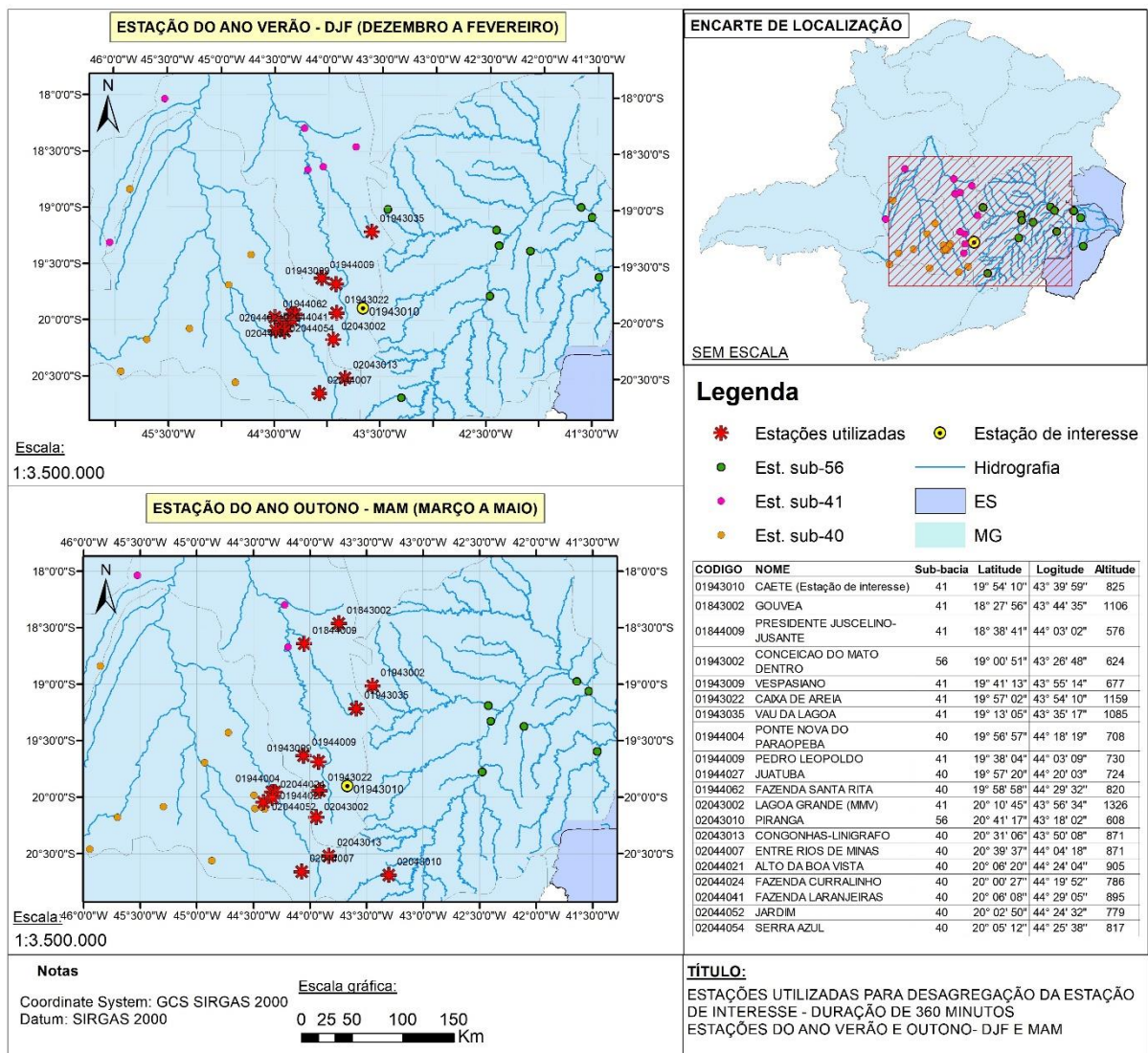
DJF		MAM		JJA		SON	
Estação	Prob.	Estação	Prob.	Estação	Prob.	Estação	Prob.
01943022	0,744	01943035	0,514	02043013	0,822	01843002	0,782
01943009	0,722	01943022	0,508	01943035	0,818	01943035	0,753
01944004	0,696	02043013	0,503	01843002	0,817	01943002	0,747
01944009	0,694	01943009	0,501	01943002	0,806	01844009	0,747
01944027	0,691	02043002	0,500	01943009	0,804	01844001	0,742
02044024	0,687	01943002	0,493	02044007	0,795	02043013	0,739
02043002	0,669	01843002	0,493	01943022	0,792	01844010	0,732
01944062	0,666	01944009	0,482	01944009	0,791	02043010	0,730
02044052	0,665	02043010	0,480	02043010	0,787	02043002	0,724
02044054	0,658	02044007	0,474	01844009	0,785	02044007	0,723
02044021	0,651	01844009	0,473	02043002	0,774	01943009	0,717
02043013	0,646	01844010	0,462	01844010	0,765	01943022	0,715
02044041	0,638	01844001	0,459	01844001	0,762	01944009	0,706
01943035	0,616	01944004	0,447	02044024	0,756	01944004	0,675
02044007	0,579	02044024	0,443	01944004	0,753	02044024	0,673
02043010	0,557	01944027	0,443	01944027	0,750	01944027	0,672
01943002	0,548	02044021	0,434	02044054	0,745	02044021	0,671
01944049	0,544	02044052	0,432	02044052	0,744	02044052	0,668
01944021	0,539	02044054	0,430	02044021	0,743	02044054	0,667
01942031	0,522	02044041	0,422	01944062	0,735	02044041	0,663
02044042	0,496	01944062	0,419	02044041	0,731	01944062	0,657
02045013	0,477	01845021	0,398	01944049	0,697	01944049	0,649
01942030	0,434	01944049	0,398	02044042	0,681	01942032	0,648
01844009	0,427	01942032	0,388	01944021	0,644	01942030	0,641
01844010	0,421	02044042	0,386	01942031	0,614	02044042	0,639
01942032	0,413	01942030	0,380	01942032	0,613	01942031	0,635
01942008	0,407	01942031	0,374	01845021	0,608	01944021	0,622
01843002	0,391	01944021	0,364	01942030	0,608	01942008	0,606
02045002	0,386	01942008	0,345	02045013	0,580	01845021	0,591
02045012	0,356	01841011	0,325	01942008	0,562	02045013	0,583
01844001	0,334	01845004	0,320	01845004	0,502	01841011	0,566
01845004	0,328	02045013	0,317	02045002	0,496	01941005	0,552
01946009	0,316	01941005	0,309	02045012	0,484	02045002	0,541
01845021	0,313	02045002	0,274	01841011	0,469	01845004	0,540
01941006	0,306	02045012	0,270	01941005	0,446	01941006	0,535
01841011	0,299	01946009	0,265	01946009	0,438	02045012	0,517
01941005	0,294	01941012	0,264	01941006	0,412	01946009	0,511
01941012	0,283	01941006	0,263	01941012	0,399	01941012	0,484
01940020	0,244	01940009	0,235	01940020	0,326	01940009	0,469
01940009	0,224	01940020	0,173	01940009	0,313	01940020	0,445



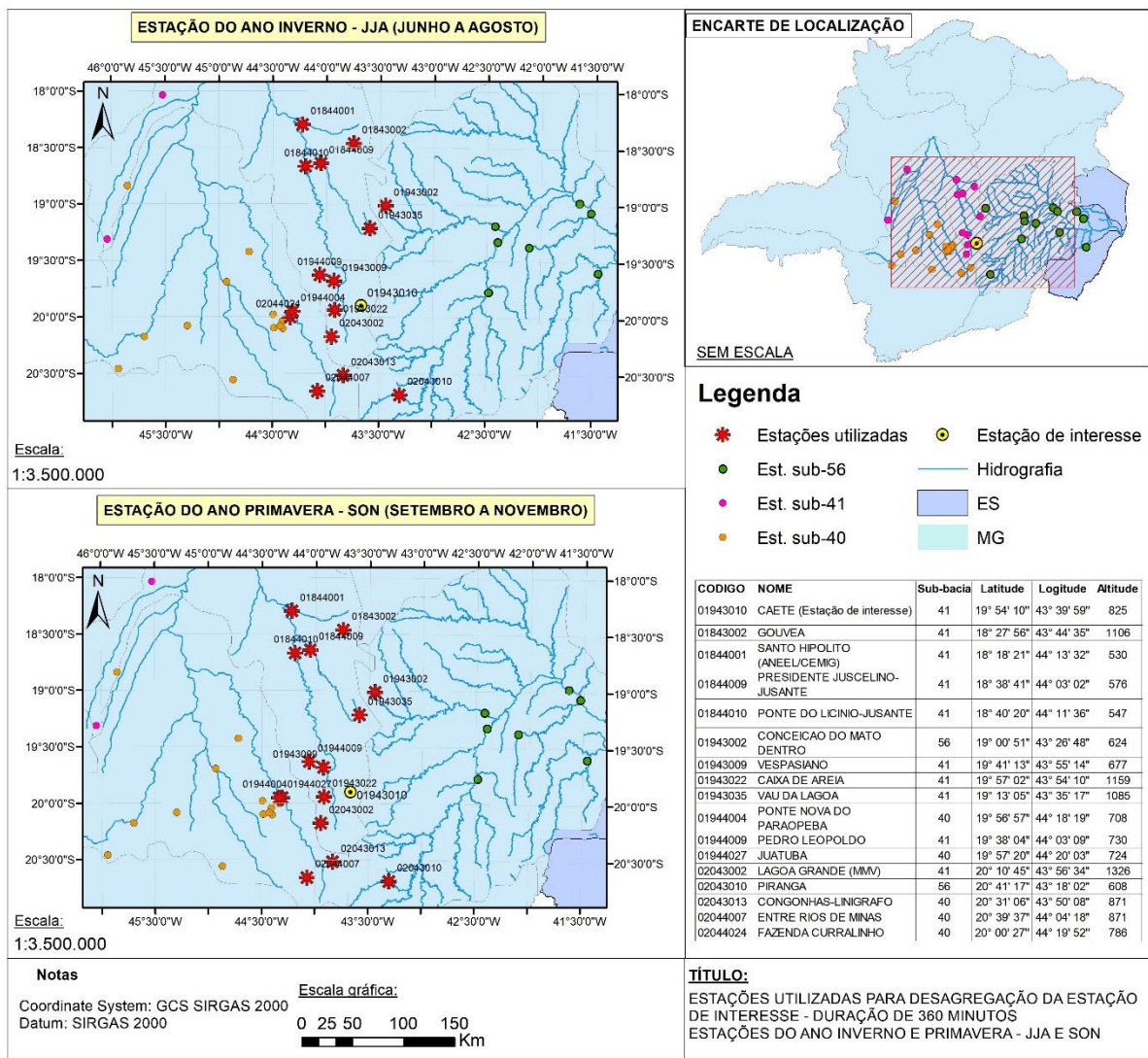
**Figura I.1 – Estações utilizadas para desagregação – Duração 180 minutos – Estações do ano Verão e Outono**



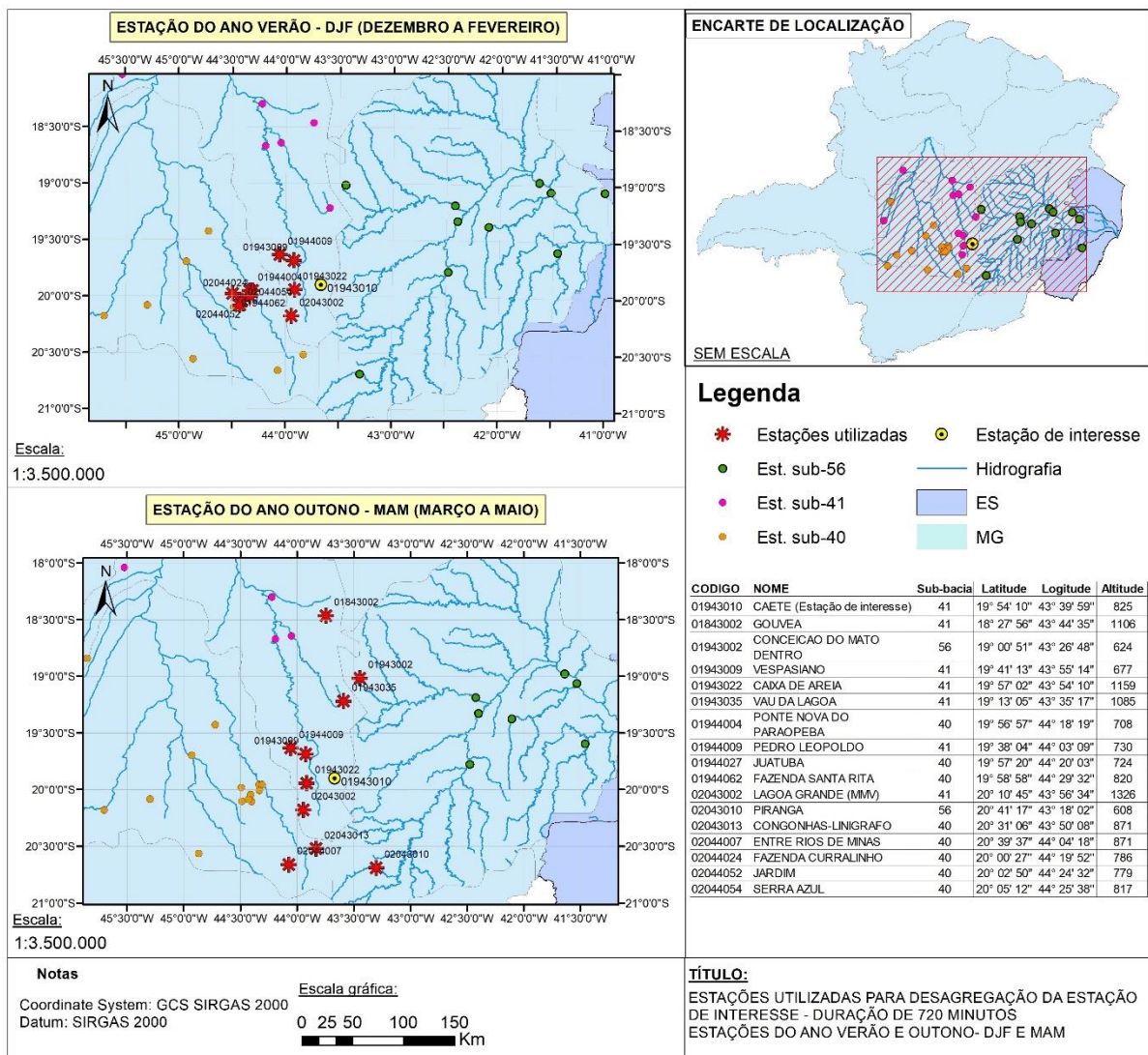
**Figura I.2 – Estações utilizadas para desagregação – Duração 180 minutos – Estações do ano Inverno e Primavera**



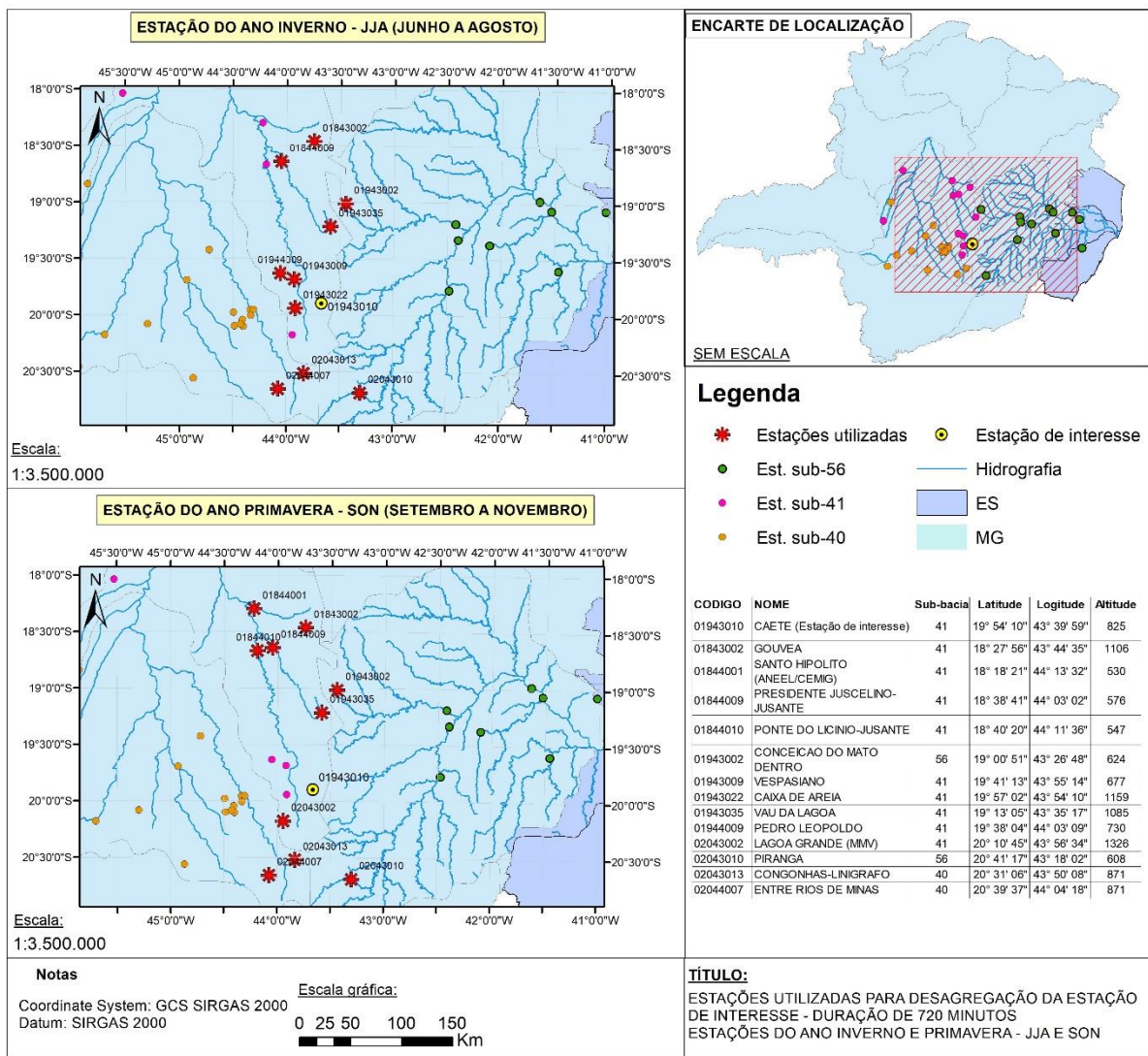
**Figura I.3 – Estações utilizadas para desagregação – Duração 360 minutos – Estações do ano Verão e Outono**



**Figura I.4 – Estações utilizadas para desagregação – Duração 360 minutos – Estações do ano Inverno e Primavera**



**Figura I.5** – Estações utilizadas para desagregação – Duração 720 minutos – Estações do ano Verão e Outono



**Figura I.6 – Estações utilizadas para desagregação – Duração 720 minutos – Estações do ano Inverno e Primavera**