

MÁRCIO JOSÉ CATALUNHA

AVALIAÇÃO DO AJUSTAMENTO DE FUNÇÕES DENSIDADE DE
PROBABILIDADE A SÉRIES DE PRECIPITAÇÃO PLUVIAL NO ESTADO DE
MINAS GERAIS

Tese apresentada à Universidade
Federal de Viçosa, como parte das
exigências do Programa de Pós-
Graduação em Meteorologia
Agrícola, para obtenção do título de
Magister Scientiae.

VIÇOSA
MINAS GERAIS - BRASIL
2000

Aos meus pais Manoel e Izabel.

Aos meus irmãos Marcelo, Cristina e Izaira.

Aos meus amigos.

AGRADECIMENTO

A Deus, pela sua infinita misericórdia e providência.

A meus pais, Manoel e Izabel, que sempre tiveram muita perseverança, domínio próprio e fé a me ensinar em todos os caminhos da vida. Meus queridos irmãos, Marcelo, Cristina e Izaíra a quem devo o incentivo e o companherismo de tantos momentos.

Ao amigo e orientador, Prof. G. C. Sedyama, que sabiamente orientou minha formação pessoal e profissional neste período.

Aos conselheiros Profs Aristides Ribeiro, Brauliro G. Leal e Carlos P. B. Soares, sempre dispostos a atender ao auxílio solicitado.

Ao Prof. Demetrius David Silva e Prof. Sérgio Zolnier, que acompanharam e incentivaram nosso trabalho.

Ao Eng. Humberto P. Euclides, que nos auxiliou em muito na coleta e tratamento dos dados para a pesquisa.

A diretoria da CAPES pelo auxílio financeiro, sem o qual nada teria sido realizado.

Aos demais professores, funcionários e alunos do Departamento de Engenharia Agrícola da UFV, pelo apoio e amizade.

BIOGRAFIA

MÁRCIO JOSÉ CATALUNHA, filho de Manoel Catalunha e Maria Izabel Dutra Catalunha, nasceu no dia 11 de setembro de 1972, na cidade de Governador Valadares, MG.

Em 1990, concluiu o curso colegial no Colégio Tiradentes da Polícia Militar, em Governador Valadares, MG.

Em julho de 1998, graduou-se em Engenharia Elétrica pela Universidade Vale do Rio Doce, Governador Valadares, MG.

Em agosto de 1998, iniciou o curso de Mestrado em Meteorologia Agrícola na Universidade Federal de Viçosa, Viçosa, MG.

CONTEÚDO

	Página
RESUMO.....	vii
ABSTRACT	ix
1. INTRODUÇÃO	1
2. OBJETIVOS	3
3. REVISÃO DE LITERATURA	4
3.1. Precipitação	4
3.2. Análise de homogeneidade	6
3.3. Distribuições de probabilidades	7
3.3.1. Distribuição exponencial	8
3.3.2. Distribuição gama.....	8
3.3.2.1. Método da máxima verossimilhança.....	9
3.3.2.2. Método de Greenwood e Durand (1960)	11
3.3.2.3. Função gama.....	12
3.3.2.4. Aproximação da distribuição gama por expansão em série.....	13
3.3.3. Distribuição log-normal.....	14
3.3.4. Distribuição normal.....	15
3.3.4.1. Aproximação da distribuição normal por expansão polinomial .	16
3.3.5. Distribuição Weibull.....	17
3.3.5.1. Método de máxima verossimilhança.....	18
3.4. Testes de aderência	20
3.4.1. Teste de aderência por χ^2 (qui-quadrado).....	20

3.4.2. Teste de aderência por Kolmogorov-Smirnov	23
3.5. Análise de agrupamento	24
3.5.1. Medidas de dissimilaridade	24
3.5.2. Métodos de agrupamento	25
4. MATERIAL E MÉTODOS	27
4.1. Seleção das estações	27
4.2. Totais de precipitação	28
4.3. Distribuições de probabilidade	30
4.3.1. Distribuição exponencial	31
4.3.2. Distribuição gama	31
4.3.3. Distribuição log-normal	32
4.3.4. Distribuição normal	32
4.3.5. Distribuição Weibull	32
4.4. Testes de aderência	33
4.5. Análise de agrupamento	33
5. RESULTADO E DISCUSSÃO	35
5.1. Seleção das estações	35
5.2. Totais médios de precipitação	37
5.3. Número médio de dias com chuva	39
5.4. Distribuições de Probabilidades	46
5.4.1. Distribuição exponencial	46
5.4.2. Distribuição gama	47
5.4.3. Distribuição log-normal	49
5.4.4. Distribuição normal	49
5.4.5. Distribuição Weibull	50
5.5. Testes de aderência	51
5.6. Análise de agrupamento	59
6. CONCLUSÕES	65
REFERÊNCIAS BIBLIOGRÁFICAS	67
APÊNDICE	70

RESUMO

CATALUNHA, Márcio José, M.S., Universidade Federal de Viçosa, agosto de 2000. **Avaliação do ajustamento de funções densidade de probabilidade a séries de precipitação pluvial no Estado de Minas Gerais.** Orientador: Gilberto C. Sedyama. Conselheiros: Brauliro Gonçalves Leal, Aristides Ribeiro e Carlos Pedro Boechat Soares.

O presente trabalho teve como objetivos específicos: testar a distribuição de probabilidade que melhor se ajusta à freqüência das precipitações; analisar os dados de precipitação, considerando sua variabilidade espacial e temporal, nos períodos anual, chuvoso, seco e mensal, bem como o número médio de dias com chuva nos mesmos períodos, estabelecendo valores médios para todo o estado; e identificar, por intermédio de técnicas de análise de agrupamento, regiões com precipitação homogênea no Estado de Minas Gerais. Foram analisadas as distribuições de probabilidade exponencial, gama, log-normal (a dois e três parâmetros), normal e Weibull. Os testes, não-paramétricos, de qui-quadrado, a 5% de significância, e de Kolmogorov-Smirnov, a 20% de significância, foram utilizados para verificar a aderência das probabilidades estimadas aos dados observados. Foram considerados, para fins de análise espacial e temporal, o total médio e número médio de dias com chuva, para os períodos anual, chuvoso, seco e mensal de precipitação. Como período chuvoso, consideraram-se os meses de janeiro, fevereiro, março, outubro, novembro e dezembro, e como período seco, os meses de abril, maio, junho, julho, agosto e setembro. Foram considerados

dias com chuva aqueles em que se observou precipitação maior ou igual a 0,1 mm. Foram feitas combinações entre os métodos de agrupamento (encadeamento simples e encadeamento completo) e as medidas de dissimilaridade (distância euclidiana, distância euclidiana média e distância de Mahalanobis), aplicados aos dados de precipitação, considerando também a otimização pelo método de Toucher. Para os valores diários de precipitação, observou-se a superioridade do ajustamento da distribuição Weibull, com exceção dos decêndios do período seco, em que predomina a distribuição exponencial. No caso dos valores totais de precipitação para o período seco, predominou a distribuição exponencial; no período chuvoso, houve uma variação entre as distribuições Weibull, exponencial, gama e normal, nesta ordem; esta última aparece somente em dois meses. A geoespacialização dos dados apresentou distinção nítida entre o período chuvoso e o seco, conforme os índices obtidos. A análise de agrupamento mostrou ser uma ferramenta importante para definir a existência de estações pluviométricas com totais médios diferentes, evitando a generalização de informações por região.

ABSTRACT

CATALUNHA, Márcio José, M.S., Universidade Federal de Viçosa, august of 2000. **Evaluation of the adjustment of probability density functions for series of pluvial precipitation of the State of Minas Gerais.** Adviser: Gilberto C. Sedyama. Committee Members: Brauliro Gonçalves Leal, Aristides Ribeiro and Carlos Pedro Boechat Soares.

This work was carried out: a) to test the probability density function that fits better to the observed precipitation frequency; b) to analyze the precipitation data, considering its time and space variability, for annual, monthly rainy, and dry periods, including the average number of days with rainfall in the annual and monthly basis. The clustering analysis technique was also used to identify areas with homogeneous precipitation in the State of Minas Gerais. The Gamma, exponential, normal, Weibull, and log-normal probability distributions (with two and three parameters) were also analyzed. The non-parametric test of chi-square, at 5% of significance, and the Kolmogorov-Smirnov test, at 20% of significance were used to verify the goodness of fit between estimated and observed rainfall data. Days with rainfall larger than 0.1 mm were considered a rainy day. Combinations among the clustering methods (simple linkage and complete linkage) and the dissimilarity measures (Euclidean distance, average Euclidean distance, and Mahalanobis distance) on precipitation data were also analyzed, using the Toucher method for optimization. For the daily precipitation values, the superiority of Weibull distribution function was demonstrated, except for ten dry days period, in which the exponential probability distribution function prevailed. For accumulated total precipitation for dry period, the exponential

probability distribution prevailed. In the rainy period, the best probabilities distributions were the Weibull distribution, exponential, gamma, and normal, in this order. The later only in two months of the year. The precipitation data geoespatialization showed clearly the distinction between rainy days, and dry periods. The clustering analysis of the rainfall data showed to be an important tool to define the existence of a particular spatial variability in the rainfall regime in the State of Minas Gerais.

1. INTRODUÇÃO

O clima é um dos principais fatores que afetam a produtividade agrícola, devido às oscilações interanuais dos seus elementos. A importância de cada elemento climático, para a produtividade agrícola, está condicionada à necessidade específica da cultura e a fatores geográficos.

Em particular, para as culturas tropicais, a precipitação pode ser considerada o elemento que mais condiciona o resultado de uma safra. Para estudos hidrológicos, no Brasil, a chuva, ou precipitação pluvial, é, evidentemente, o tipo mais importante, embora existam outros tipos de precipitação (garoa, neve e granizo).

Apesar de a água ser uma substância abundante, nos dias atuais acentua-se sua carência, principalmente no que se refere à água de boa qualidade. Além disso, a elevação do custo da energia também tem afetado o custo do seu deslocamento até o local de utilização. Assim, torna-se fundamental o conhecimento da quantidade e época de aplicação d'água em uma cultura, evitando seu déficit ou minimizando seu desperdício.

O sucesso da produção agrícola, principalmente em áreas não-irrigadas, é altamente dependente das três principais características do regime pluviométrico: quantidade, frequência e intensidade. O conhecimento dessas características é componente básico para o planejamento de atividades agrícolas, como definição de datas mais apropriadas para o preparo do solo e plantio, definição de macrorregiões com potencial agrícola e elaboração de projetos de irrigação suplementar. No entanto, a exploração de determinadas

culturas é mais dependente da distribuição da precipitação do que do volume total precipitado anualmente.

O aproveitamento adequado dos recursos hídricos requer o uso de técnicas de planejamento que se baseiam na estimativa das probabilidades associadas a certas variáveis hidrológicas, no caso a precipitação. Nesses trabalhos, o principal interesse é prever, com base em funções de densidade de probabilidade aplicadas aos dados observados, as precipitações capazes de ocorrer em uma certa localidade.

O uso de funções densidade de probabilidade está diretamente ligado à natureza dos dados a que ela se relaciona. Algumas têm boa capacidade de estimação para pequeno número de dados, outras requerem grande série de observações. Devido ao número de parâmetros de sua equação, algumas podem assumir diferentes formas, enquadrando-se em um número maior de situações, ou seja, são mais flexíveis. Desde que respeitado o aspecto da representatividade dos dados, as estimativas dos seus parâmetros para uma determinada região podem ser estabelecidas como de uso geral, sem prejuízo da precisão na estimação da probabilidade.

Em todos os setores da economia, é possível encontrar exemplos nos quais os conhecimentos sobre as distribuições de chuvas, no espaço e no tempo, desempenham papel relevante no processo de tomada de decisão, quer seja em nível operacional, quer seja em nível de planejamento. Pode-se notar que o amplo conhecimento do comportamento da precipitação em uma região é de extrema importância para diversas atividades.

O presente trabalho não pretendeu avaliar a precipitação com base em nenhuma outra variável, como relevo, vegetação, solo, temperatura ou atividade agroindustrial, devido à complexidade do tratamento dessas interações.

Com base na necessidade e importância dos dados pluviométricos, o presente trabalho teve por objetivo geral avaliar estatisticamente a distribuição de probabilidade que melhor se ajusta à natureza dos dados pluviométricos de Minas Gerais, bem como mostrar a variação espacial e temporal da precipitação no Estado de Minas Gerais e realizar um estudo com técnicas de análise de agrupamento.

2. OBJETIVOS

Este trabalho teve como objetivos:

- Testar o ajuste de diferentes funções densidade de probabilidade aos dados pluviométricos do Estado de Minas Gerais.
- Analisar os dados de precipitação pluvial, considerando sua variabilidade espacial e temporal, nos períodos anual, chuvoso, seco, mensal, considerando também o número médio de dias com chuva nos períodos anual e mensal e estabelecendo valores médios para todo o estado.
- Identificar, por meio de técnicas de análise de agrupamento, regiões com precipitações homogêneas no Estado de Minas Gerais.

3. REVISÃO DE LITERATURA

3.1. Precipitação

A precipitação é definida como toda a água, na forma líquida ou sólida, que cai da atmosfera em direção ao solo. Esta é a maior classe de hidrometeoros, devendo ser distinguida de outras classes, como a de nuvens, nevoeiros, orvalho, geada, etc. Vale ressaltar que os dois últimos, orvalho e geada, formam-se junto ao solo, não devendo, portanto, ser confundidos com precipitação. Em escala macroscópica, a precipitação pluvial se origina de nuvens formadas pelo resfriamento de uma massa de ar, ocasionado pela sua expansão ao se elevar na atmosfera (HUSCHKE, 1970).

A quantidade de precipitação que incide numa bacia hidrográfica durante o ano é um dado determinante para quantificar, entre outras, a disponibilidade de água para o abastecimento doméstico, industrial e irrigação.

A distribuição da chuva, no espaço e no tempo, é de fundamental importância no dimensionamento de obras hidráulicas, como drenos, barragens, diques de proteção contra cheias, bueiros de estradas, etc. O dimensionamento destas obras passa também obrigatoriamente pelo estudo das chuvas intensas ocorridas no local (BELTRAME, 1979).

As chuvas ácidas estão se transformando num problema para a sobrevivência da humanidade. Um número elevado de lagos, na Europa, no Canadá e em países do Terceiro Mundo, está se tornando ácidos, aumentando

dessa forma a concentração de alumínio livre, que é letal para a vida aquática (HAAG, 1985).

A chuva é o principal agente que intervém no processo erosivo do solo. O potencial erosivo da chuva, para cada localidade, é decorrente das características individuais de cada chuva. Pesquisas realizadas por FRANCO (1986), no Paraguai, mostraram também que a erodibilidade do solo poderia ser indicada pelo nível de precipitação e pelo tipo de solo.

Para se obter o número de dias trabalháveis com tratores agrícolas, é necessário ter conhecimento sobre a umidade do solo e a quantidade de precipitação ocorrida (ASSIS, 1988).

TOMMASELLI (1992) afirma que, com a irrigação, pode-se plantar, algumas culturas, em qualquer época, mas que a adoção dos períodos mais favoráveis à precipitação reduz muito os custos de produção, que vêm ganhando cada vez mais importância no mercado econômico.

Segundo CUPOLILLO (1997), o El Niño e a La Niña são eventos complementares constituintes de uma variação positiva e negativa conhecida como Índice de Oscilação Sul (IOS). Estes eventos estão claramente relacionados aos eventos climáticos sobre toda a massa global, provocando modificações de temperatura e precipitação em todas as zonas de sua ocorrência.

Vários critérios têm sido propostos para caracterizar um dia seco. CHATFIELD (1940) recomenda o valor de 1,0 mm de chuva diária, abaixo do qual se considera como dia seco. Gates e Tong (1976), citados por ALMEIDA (1995), consideraram a precipitação diária até 0,1 mm como dia seco. O critério adotado por WOLF (1962) considera dia seco aquele em que a evapotranspiração real é maior que a chuva ocorrida nesse mesmo dia. Stern (1980), citado por ALMEIDA (1995), define dia seco como aquele em que o total de precipitação for menor ou igual a 0,3 mm. PAIS DE MELO (1989) considerou dias chuvosos aqueles em que o total de precipitação foi maior ou igual a 0,1 mm.

Em termos ambientais, o conhecimento do número médio de dias com chuva é importante, segundo ASSAD (1994), para:

- quantificar os dias de chuva no período chuvoso, identificando, assim, os períodos mais favoráveis ao aparecimento dos problemas de erosão;

- quantificar os dias de chuva no período seco, identificando, dessa forma as áreas mais susceptíveis às queimadas; e
- quantificar os dias de chuva no início e fim do período chuvoso, podendo estabelecer, assim, os dias disponíveis de trabalho para a agricultura, envolvendo as práticas de plantio e colheita.

3.2. Análise de homogeneidade

Uma série numérica ordenada no tempo, representando as variações de um elemento climatológico, é dita “homogênea” se estas variações são causadas pelas variações do tempo e do clima. Conseqüentemente, se influências não-naturais, como erros instrumentais, de observadores, etc, atuam sobre as medições, a série é dita “não-homogênea”.

Não é tarefa fácil constatar a “não-homogeneidade” de uma série, principalmente porque as dificuldades variam de um elemento para outro. Por outro lado, não é possível distinguir entre variações naturais e artificiais de um elemento climático se apenas se dispõe de observações de um único local. Isso significa que é geralmente difícil decidir se uma série é “absolutamente homogênea”.

A *World Meteorological Organization* (WMO), em sua nota técnica, nº 81, recomenda um teste não-paramétrico para testar a homogeneidade de uma série, o qual é sensível a desvios da média e tendências. Este teste é denominado *run test*, ou teste de seqüência, e consiste na contagem do número de ocorrências de valores observados situados acima e abaixo da mediana, testando o resultado da contagem, mediante o uso de uma tabela de distribuição (Tabela 1).

Tabela 1 - Distribuição do número de seqüências para o *run test*

NA	Limite a 0,10	Limite a 0,90
10	8	13
11	9	14
12	9	16
13	10	17
14	11	18
15	12	19
16	13	20
17	14	21
18	15	22
19	16	23
20	16	25
25	22	30
30	26	36
35	31	41
40	35	47
45	40	52
50	45	57

NA é igual à metade do número de observações; 0,10 e 0,90 são os limites críticos para tendência e oscilação, respectivamente.

Assim, está claro que excesso de seqüências seria uma indicação de oscilação. Poucas seqüências indicam tendências ou afastamento da mediana. Um valor dentro do intervalo da tabela não é significativamente diferente do esperado, e a série é dita “homogênea”.

3.3. Distribuições de probabilidades

Cada distribuição de freqüência apresenta uma certa forma, esta forma pode ser aproximada através da utilização de equações de densidade probabilidade com alguns parâmetros extraídos da amostra em questão. A utilização ou não de uma distribuição reside na capacidade da mesma em estimar os dados observados, e esta capacidade é medida através de testes de aderência. A seguir é apresentada as distribuições utilizadas nesta análise.

3.3.1. Distribuição exponencial

A distribuição exponencial é geralmente aplicada a dados com forte assimetria, ou seja, apresentando uma forma de "J" invertido. Sua função densidade de probabilidade é assim descrita (KITE, 1978):

$$f(x) = \frac{\exp\left(-\frac{x}{\bar{x}}\right)}{\bar{x}} \quad \text{para } x > 0 \quad 1$$
$$f(x) = 0 \quad \text{para } x < 0$$

sua função de distribuição acumulada é do tipo:

$$F(x) = \int_0^x f(x) = 1 - \exp\left(-\frac{x}{\bar{x}}\right) \quad 2$$

O único parâmetro da distribuição, como visto, é a média.

3.3.2. Distribuição gama

Stern e Coe (1982), citados por ALMEIDA (1995), afirmam que o ajuste de modelos probabilísticos aos dados diários de chuva, além de fornecer um resumo sucinto destes dados, representa uma técnica eficiente para a análise dessas informações. Os autores testaram a distribuição gama para modelar as quantidades diárias de chuva, em localidades da Jordânia, Nigéria, de Botswana e Sri Lanka, tendo obtido resultados satisfatórios. FARIA (1998) utilizou a distribuição gama para estimar a precipitação dependente em nível de 75% de probabilidade, obtendo boa aderência com o teste de Kolmogorov-Smirnov.

Se x for uma variável aleatória contínua, tal que $(0 < x < \infty)$, com distribuição gama de parâmetros $\alpha > 0$ e $\beta > 0$, então a sua função densidade de probabilidade é definida como:

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad , 0 < x < \infty \quad 3$$

Essa função densidade de probabilidade descreve uma família de curvas que depende dos parâmetros α e β .

Sendo $F(x)$ a probabilidade de ocorrência de um evento menor ou igual a x , pode-se escrever que a função de distribuição acumulada de probabilidade é representada pela função gama incompleta, segundo THOM (1958):

$$F(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^x u^{\alpha-1} e^{-\frac{u}{\beta}} du \quad 4$$

em que,

- $F(x)$ = probabilidade de ocorrer um valor menor ou igual a x ;
- x = variável aleatória contínua;
- $\Gamma(\alpha)$ = função gama do parâmetro alfa;
- α = parâmetro de forma da variável aleatória x ;
- β = parâmetro de escala da variável aleatória x ; e
- e = base do logaritmo neperiano (2,718281828...).

Algumas formas de estimar os parâmetros da distribuição gama foram desenvolvidas, contribuindo, junto com a sua flexibilidade de formas, para sua utilização em diversas áreas (HAAN, 1977). Os principais métodos para estimar seus parâmetros são:

- a) método de máxima verossimilhança; e
- b) método de Greenwood e Durand (1960).

3.3.2.1. Método da máxima verossimilhança

O método da máxima verossimilhança desenvolvido por Fisher (1941), citado por COUTO (1980), é um processo geral de obtenção de estimadores. A função de verossimilhança é dada por.

$$L(x_1, x_2, \dots, x_n; \alpha, \beta) = \prod_{i=1}^n \frac{1}{\beta^\alpha \Gamma(\alpha)} (x_i)^{\alpha-1} e^{-\frac{x_i}{\beta}} \quad 5$$

Aplicando logaritmo, obtém-se

$$\ln(x_1, x_2, \dots, x_n; \alpha, \beta) = -n\alpha \ln(\beta) - \ln[\Gamma(\alpha)] + (\alpha - 1) \ln(\sum \ln x_i) - \frac{\sum x_i}{\beta} \quad 6$$

Tomando-se a derivada parcial, igualando-se a zero e simplificando, obtém-se

$$\begin{cases} -n \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - n \ln \hat{\beta} + \sum_{i=1}^n \ln x_i = 0 \\ \hat{\beta} = \frac{\bar{x}}{\hat{\alpha}} \end{cases} \quad 7$$

A equação pode ser representada por

$$\ln \hat{\alpha} - \psi(\hat{\alpha}) = \ln \bar{x} - \frac{1}{n} \sum_{i=1}^n \ln x_i \quad 8$$

Fazendo-se

$$A = \ln \bar{x} - \frac{1}{n} \sum_{i=1}^n \ln x_i \quad 9$$

em que,

- n = número de dados da série;
- x_i = valor da variável no período i da série;
- \bar{x} = valor médio da variável x; e
- i = i-ésimo valor da variável x.

Obtém-se

$$\ln \hat{\alpha} - \psi(\hat{\alpha}) = A \quad 10$$

A dificuldade do método reside na resolução da Equação 10, de onde obtém-se o estimador $\hat{\alpha}$, isto é, é uma equação implícita em $\hat{\alpha}$. A função digama aparece tabulada em algumas publicações, como em ABRAMOWITZ e STEGUN (1972). A função digama, segundo THOM (1958), pode ser obtida através do desenvolvimento de

$$\psi(\alpha) = \ln(\alpha) - \frac{1}{2\alpha} - \sum_{k=1}^m \frac{B_{2k}}{2k\alpha^{2k}} \quad 11$$

em que B_k são os números de Bernoulli ($B_2=1/6$, $B_4=-1/30, \dots$).

Simplificando a expressão de $\psi(\hat{\alpha})$ da Equação 11, um estimador aproximado do parâmetro α pode ser obtido através da resolução da seguinte equação:

$$12A\hat{\alpha}^2 - 6\hat{\alpha} - 1 = 0 \quad 12$$

Então, para satisfazer a condição $\alpha > 0$ (por definição), a solução que interessa será

$$\hat{\alpha} = \frac{1 + \sqrt{1 + 4A/3}}{4A} \quad 13$$

em que,

$$A = \ln(\bar{x}) - x_g \quad 14$$

O estimador do parâmetro β , como descrito anteriormente, é

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}} \quad 15$$

sendo \bar{x} a média aritmética e x_g a média geométrica das observações.

3.3.2.2. Método de Greenwood e Durand (1960)

O método proposto por Greenwood e Durand, em 1960, utiliza duas equações para intervalos diferentes dos valores de A (Equação 14) (Greenwood e Durand, 1960). Para o primeiro intervalo, tem-se que α é calculado por

$$\alpha = \frac{0,5000876 + 0,1648852A - 0,054427A^2}{A} \quad 16$$

quando $0 \leq A \leq 0,5772$, por

$$\alpha = \frac{8,898919 + 9,05885A + 0,9775373A^2}{A(17,79728 + 11,968477A + A^2)} \quad 17$$

quando $0,5772 < A \leq 17$, em que o parâmetro β é calculado pela Equação 15.

3.3.2.3. Função gama

Como pode ser notado na equação de definição da distribuição gama, Equação 4, o cálculo do valor de gama é um parâmetro importante para a distribuição.

A função gama é obtida a partir da transformada de Laplace da função t^α para $\alpha > -1$ (SEDIYAMA, 1999), em que α é uma variável aleatória:

$$L\{t^\alpha\} = \int_0^\infty t^\alpha e^{-u} dt, \alpha > -1 \quad 18$$

Fazendo as devidas substituições e simplificações, tem-se

$$L\{t^\alpha\} = \frac{1}{s^{\alpha+1}} \int u^\alpha e^{-u} du, (\alpha > -1) \quad 19$$

Substituindo α por $\alpha-1$, convenientemente a função toma a forma da função gama:

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du, \alpha \geq 0 \quad 20$$

O processo de integração torna-se difícil, tendo em vista que a distribuição não tem limite superior. Surgem então funções aproximadoras para este cálculo. Em uma de suas formas, a função gama pode ser escrita como:

$$\Gamma(\alpha) = \Gamma(n+N) = \left[\prod_{k=1}^{N-1} (n+N-k) \right] n\Gamma(n), \text{ com } \alpha > 0 \quad 21$$

em que,

- n = parte fracionária de α ;
 - N = parte inteira de α ; e
 - k = 1,2,3...N.
- E considerando

$$\Gamma(n) = \frac{\Gamma(n+1)}{n}, \text{ com } 0 < n \leq 1 \quad 22$$

ABRAMOWITZ e STEGUN (1972) definem um polinômio aproximador para o cálculo da função gama no intervalo de $0 \leq n \leq 1$, que pode ser estabelecido da seguinte forma:

$$\Gamma(n+1) = 1 + a_1n + a_2n^2 + a_3n^3 + a_4n^4 + a_5n^5 + a_6n^6 + a_7n^7 + a_8n^8 \quad 23$$

em que,

$$a_{1=} -0,577191652; \quad a_{5=} -0,756704078;$$

$$\begin{array}{ll}
a_2= & +0,988205891; & a_6= & +0,482199394; \\
a_3= & -0,897056937; & a_7= & -0,193527818; \\
a_4= & +0,918206857; & a_8= & +0,035868343.
\end{array}$$

Segundo os autores citados, a função gama também pode ser aproximada por duas outras funções, como a fórmula de Stirling's

$$\Gamma(\alpha) = e^{-\alpha} \alpha^{\alpha-0.5} \sqrt{2\pi} \left[1 + \frac{1}{12\alpha} + \frac{1}{288\alpha^2} - \frac{139}{51840\alpha^3} - \frac{571}{2488320\alpha^4} + \dots \right] \quad 24$$

com $\alpha > 0$ e a fórmula assintótica

$$\ln \Gamma(\alpha) = (\alpha - 0.5) \ln \alpha - \alpha + 0.5 \ln(2\pi) + \frac{1}{12\alpha} - \frac{1}{360\alpha^3} + \frac{1}{1260\alpha^5} - \frac{1}{1680\alpha^7} + \dots \quad 25$$

com $\alpha > 0$.

ASSIS et al. (1996) citam uma outra relação para aproximação da função gama

$$\Gamma(\alpha) = \sqrt{\frac{2\pi}{\alpha}} e^{\alpha[\ln \alpha - f(\alpha)]}, \text{ com } \alpha > 0 \quad 26$$

$$f(x) = 1 - \frac{1}{12\alpha^2} + \frac{1}{360\alpha^4} - \frac{1}{1260\alpha^6} \quad 27$$

3.3.2.4. Aproximação da distribuição gama por expansão em série

A função de distribuição acumulada da distribuição de probabilidade gama possui integral que pode ser resolvida por métodos numéricos ou pelo desenvolvimento em série de uma expressão exponencial. Para isso, utilizando-se de transformação de variáveis, tal que $t = u/\beta \rightarrow du = \beta dt$, não mudando o limite inferior zero e substituindo x por t , e após as simplificações adequadas, tem-se

$$F(t) = \frac{1}{\Gamma(\alpha)} \int_0^t t^{\alpha-1} e^{-t} dt \quad 28$$

Considerando o desenvolvimento em série de e^{-t} , tem-se

$$e^{-t} = 1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + \dots \quad 29$$

e, substituindo na Equação 28, pode-se escrever:

$$F(t) = \frac{1}{\Gamma(\alpha)} \int_0^t t^{\alpha-1} \left(1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + \dots \right) dt \quad 30$$

Multiplicando-se e dividindo-se a equação anterior por e^{-t} :

$$F(t) = \frac{t^\alpha}{\Gamma(\alpha)e^t} \left[\frac{1}{\alpha} - \frac{t}{\alpha+1} + \frac{t^2}{(\alpha+2)2!} - \frac{t^3}{(\alpha+3)3!} + \dots \right] \quad 31$$

Reunindo, do produto resultante, os termos em t , t^2 , t^3 ,... e os demais termos correspondentes aos diversos expoentes, tem-se, segundo ASSIS et al. (1996):

$$F(t) = \frac{t^\alpha}{\alpha\Gamma(\alpha)e^t} \left[1 + \frac{t}{\alpha+1} + \frac{t^2}{(\alpha+1)(\alpha+2)} + \frac{t^3}{(\alpha+1)(\alpha+2)(\alpha+3)} + \dots \right] \quad 32$$

em que, $t = x/\beta$, é uma aproximação da distribuição gama por expansão em série.

3.3.3. Distribuição log-normal

Uma outra distribuição, testada por HUF e NEIL (1959) num trabalho de comparação entre vários métodos para analisar freqüência de precipitação, é a distribuição log-normal, a qual assume que os logaritmos das variáveis aleatórias são normalmente distribuídos.

Conforme MIRSHAWKA (1971), a função densidade da distribuição log-normal a dois parâmetros e a três parâmetros são representadas pelas seguintes equações, respectivamente:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{[\ln x - \mu]^2}{2\sigma^2}\right) \quad 33$$

$$f(x) = \frac{1}{(x-a)\sigma\sqrt{2\pi}} \exp\left(-\frac{[\ln(x-a) - \mu]^2}{2\sigma^2}\right) \quad 34$$

em que,

- $f(x)$ = probabilidade da intensidade de uma variável ser menor ou igual a x ;
- e = base dos logaritmos neperianos;
- x = valor da variável aleatória;
- μ = média dos logaritmos da variável x ;
- σ = desvio-padrão dos logaritmos da variável x ; e
- a = limite inferior da amostra.

Para encontrar a probabilidade de que uma variável aleatória x tendo distribuição log-normal, assuma valores entre a e b ($a \leq x \leq b$), tem-se

$$F(a \leq x \leq b) = \int_a^b \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{[\ln(x-a) - \mu]^2}{2\sigma^2}\right) dx \quad 35$$

O valor de “ a ” pode ser zero, quando se considera a distribuição log-normal a dois parâmetros, ou um valor mínimo da série, quando se considera log-normal a três parâmetros. O valor de “ b ” pode ser o da variável aleatória, quando se considera a probabilidade cumulativa de ocorrência daquele valor (HASTINGS e PEACOCK, 1975).

3.3.4. Distribuição normal

A distribuição de probabilidade contínua mais utilizada é a distribuição normal (HASTINGS e PEACOCK, 1975), geralmente citada como curva normal ou curva de Gauss. Sua importância em análise matemática resulta do fato de que muitas técnicas estatísticas, como análise de variância, de regressão e alguns testes de hipótese, assumem ou exigem a normalidade dos dados.

A distribuição normal é uma distribuição de dois parâmetros. Sua função densidade de probabilidade tem a seguinte forma:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \text{ para } x \geq 0 \quad 36$$
$$f(x) = 0 \text{ para } x < 0$$

em que μ é a média e σ o desvio-padrão da variável aleatória.

A probabilidade de que um x assumira valores menores ou iguais quando ele tem $N(\mu, \sigma^2)$, distribuição normal com média μ e variância σ^2 , é estimada por

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \quad 37$$

Mas essa equação não pode ser resolvida analiticamente sem o uso de métodos de integração aproximada. Por esse motivo usa-se a transformação de $Z = \frac{x-\mu}{\sigma}$, a variável Z tem $N(0,1)$, distribuição normal com média zero e variância um. A variável Z é chamada variável reduzida, e a Equação 37 toma a seguinte forma:

$$F(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{Z^2}{2}\right) dz \quad 38$$

3.3.4.1. Aproximação da distribuição normal por expansão polinomial

PACITTI (1974) argumenta que, para evitar grande trabalho de computação no cálculo da função da distribuição normal padrão, usa-se um polinômio que aproxima a função com boa precisão. Este polinômio é o apresentado a seguir:

$$F(Z) \approx 1 - \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z^2}{2}\right) (a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5) \quad 39$$

para $Z \geq 0$, já que para $Z < 0$ tem-se $F(Z) = 1 - F(-Z)$

$$\begin{aligned} t &= \frac{1}{1+kZ}; \\ k &= +0,2316419; \\ a_1 &= +0,31938153; \\ a_2 &= -0,356563782; \\ a_3 &= +1,781477937; \\ a_4 &= -1,821255978; \\ a_5 &= +1,330274429; \text{ e} \\ Z &= \text{variavel reduzida ou transformada de } x. \end{aligned}$$

3.3.5. Distribuição Weibull

A distribuição de probabilidade Weibull foi proposta primeiramente por Fisher e Tippett em 1928, tendo sido desenvolvida independentemente por Walodi Weibull, físico sueco, em 1939, no estudo de resistência de materiais. Provavelmente, após a Segunda Guerra Mundial, com a ênfase no estudo da resistência dos materiais, o trabalho de Weibull se destacou e a distribuição passou a ser chamada pelo seu nome (JOHNSON e KOTZ, 1970).

Sua função de densidade de probabilidade é apresentada de diversas formas, sendo comum em alguns trabalhos apresentá-la como

$$f(x) = \frac{\gamma}{\beta} \left(\frac{x - \alpha}{\beta} \right)^{\gamma-1} \exp \left[- \left(\frac{x - \alpha}{\beta} \right)^{\gamma} \right] \text{ para } x \geq \alpha \quad 40$$

em que $f(x) = 0$ para outros intervalos, x é a variável aleatória e $\alpha \geq 0$, $\beta > 0$ e $\gamma > 0$ os parâmetros da distribuição.

Nessa forma, a distribuição é normalmente chamada de Weibull a três parâmetros. A sua função de distribuição acumulada é

$$F(x) = \int_0^{\infty} f(x) dx = 1 - \exp \left[- \left(\frac{x - \alpha}{\beta} \right)^{\gamma} \right] \quad 41$$

Uma das razões que fazem a distribuição ser aplicada em muitas situações é a íntima relação entre seus parâmetros e o comportamento da distribuição.

O parâmetro α é chamado de parâmetro de posição, pois controla a posição da curva sobre o eixo das abscissas. Quando α assume o valor zero (ou o menor valor da amostra) a distribuição inicia-se neste valor, surgindo um caso especial, sendo então chamada de Weibull a dois parâmetros.

O parâmetro de escala β controla as dimensões que a curva assume, dada uma forma constante. A medida que o valor de β aumenta, a curva se torna mais dispersa; isso pode ser facilmente notado quando a Weibull assume a forma de um sino.

A grande variedade de formas que a distribuição Weibull pode assumir é controlada pelo parâmetro γ , chamado parâmetro da forma. Quando $\gamma \leq 1$, a

distribuição assume a forma de “J” invertido. Para valores de γ entre 1 e 3,6, a distribuição torna-se positivamente assimétrica, tomando a forma de um sino, praticamente simétrica e semelhante à distribuição normal quando $\gamma = 3,6$. À medida que o valor de γ passa de 3,6 e tende ao infinito, a distribuição torna-se negativamente assimétrica, tendendo a uma forma extremamente leptocúrtica.

Algumas propriedades da Weibull a dois parâmetros são apresentadas por JOHNSON e KOTZ (1970).

Muitas maneiras de estimar os parâmetros da distribuição Weibull foram desenvolvidas, contribuindo, junto com a sua grande flexibilidade de forma, para sua utilização em diversas áreas. A maior parte dos estimadores se refere a Weibull a dois parâmetros, pois seu ajuste é mais simples e sua utilização mais freqüente.

Os principais métodos de ajuste da distribuição Weibull são (COUTO, 1980):

- máxima verossimilhança;
- momentos da distribuição;
- percentis da distribuição;
- funções lineares; e
- regressão linear ou não-linear.

Observou-se que, em geral, os estimadores de máxima verossimilhança são mais precisos, devido à sua consistência matemática, e de fácil implementação.

3.3.5.1. Método de máxima verossimilhança

COUTO (1980) apresenta os estimadores de máxima verossimilhança da Weibull a dois parâmetros. A dedução da função de máxima verossimilhança é dada a seguir.

A função densidade de probabilidade é dada por

$$f(x) = \frac{\gamma}{\beta} \left(\frac{x}{\beta}\right)^{\gamma-1} \exp\left(-\left(\frac{x}{\beta}\right)^{\gamma}\right)$$

42

Para a função de máxima verossimilhança, tem-se

$$L = \prod f(x) = \left(\frac{\gamma}{\beta}\right)^n \prod \left(\frac{x}{\beta}\right)^{\gamma-1} \exp\left(-\sum \left(\frac{x}{\beta}\right)^\gamma\right) \quad 43$$

O logaritmo da função de máxima verossimilhança será

$$\ln(L) = \ln\left(\frac{\gamma^n}{\beta^n}\right) + \ln\left(\prod x^{\gamma-1}\right) - \frac{1}{\beta^\gamma} \sum x^\gamma \quad 44$$

A primeira derivada do logaritmo da função de máxima verossimilhança em relação ao parâmetro β é

$$\frac{\partial \ln(L)}{\partial \beta} = -\frac{n\gamma}{\beta} - \left(\frac{-\gamma}{\beta^{\gamma+1}}\right) \sum x^\gamma \quad 45$$

O estimador do parâmetro β será

$$\hat{\beta} = \left(\frac{\sum x^\gamma}{n}\right)^{\frac{1}{\gamma}} \quad 46$$

A primeira derivada do logaritmo da função de máxima verossimilhança em relação ao parâmetro γ é

$$\ln(L) = n \ln(\gamma) - n\gamma \ln(L) + \gamma \sum \ln(x) - \sum \ln(x) - \sum \ln\left(\frac{x}{\beta}\right)^\gamma \quad 47$$

Reorganizando, o estimador do parâmetro γ será

$$\hat{\gamma} = \frac{n}{\frac{1}{\beta^\gamma} \sum (x^\gamma \ln(x)) - \sum \ln(x)} \quad 48$$

Portanto, o estimador dos parâmetros β e γ , para máxima verossimilhança, são, respectivamente:

$$\hat{\gamma} = n / \left[\left(\frac{1}{\hat{\beta}^{\hat{\gamma}}} \right) \sum_{i=1}^n (x_i^{\hat{\gamma}} \ln(x_i)) - \sum_{i=1}^n \ln(x_i) \right] \quad 49$$

$$\hat{\beta} = \left(\frac{\sum_{i=1}^n \ln(x_i^{\hat{\gamma}})}{n} \right)^{\frac{1}{\hat{\gamma}}} \quad 50$$

Nota-se que β é função de γ , enquanto este é função de β e de si próprio. Normalmente, a solução desse sistema é obtido pela substituição da Equação 50 na Equação 49 e transformando para

$$\frac{\sum_{i=1}^n x_i^{\hat{\gamma}} \ln(x_i)}{\sum_{i=1}^n x_i} - \frac{\sum_{i=1}^n \ln(x_i)}{n} - \frac{1}{\hat{\gamma}} = 0 \quad 51$$

Encontra-se γ resolvendo a equação anterior por um processo iterativo (em geral, o método da bissecção), sendo β encontrado a partir de γ .

3.4. Testes de aderência

Ao se ajustar uma distribuição de probabilidade a um conjunto de dados, trabalha-se com a hipótese de que a distribuição pode representar adequadamente aquele conjunto de informações. Uma maneira de comprovar esta hipótese é através de alguns testes não paramétricos, como os mencionados a seguir.

3.4.1. Teste de aderência por χ^2 (qui-quadrado)

O teste do χ^2 é aplicado para verificar o ajustamento de uma distribuição de probabilidade específica conhecida a uma amostra de dados de uma distribuição desconhecida. A hipótese de nulidade admite que a distribuição seja a especificada (gama, Weibull etc...), com os seus parâmetros estimados com base nos dados amostrais.

A hipótese de nulidade é testada fazendo-se a comparação entre as frequências observadas e as frequências teóricas, em cada classe de frequência da amostra, com a variável aleatória χ^2 dada por (CAMPOS, 1979):

$$\chi^2 = \sum_{i=1}^k \left(\frac{(F_{o_i} - F_{e_i})^2}{F_{e_i}} \right) \quad 52$$

em que, k é o número de classes, F_{o_i} a freqüência observada e F_{e_i} a freqüência esperada, de acordo com a distribuição que está sendo testada.

Esta variável qui-quadrado, χ^2 , tem uma distribuição com um único parâmetro, o grau de liberdade (GL), dado por

$$GL = k - p - 1 \quad 53$$

em que, k é o número de classes e p o número de parâmetros da distribuição sob teste.

A hipótese testada é

$$\chi_{(obs)}^2 < \chi_{(1-\alpha, k-p-1)}^2 \quad 54$$

Em que $\chi_{(obs)}^2$ é o valor de qui-quadrado observado; $\chi_{(1-\alpha, k-p-1)}^2$ é o valor tabelado de qui-quadrado com $(k-p-1)$ graus de liberdade e α é o nível de significância estabelecido.

Para uso deste teste é feita uma única restrição: que sejam reunidas em uma única classe as classes com freqüência esperada, F_{e_i} , baixas; ou seja, menores que 3 ou 5 (ASSIS et al., 1996).

Pode-se perceber, intuitivamente, que, quanto maior o valor de χ^2 , maior será a probabilidade de as freqüências observadas não serem provenientes da população teórica, admitida *a priori*.

Valores de χ^2 além do valor crítico, com grau de liberdade e nível de significância especificado, leva à rejeição da hipótese de nulidade. No entanto, se χ^2 for pequeno, ou seja, aquém do valor crítico, é porque existe concordância entre as freqüências observadas e esperadas e a amostra provém de uma população que segue a distribuição de probabilidade sob teste. Os valores críticos de χ^2 para alguns níveis de significância são descritos na Tabela 2.

Tabela 2 - Valores críticos de χ^2

Grau de Liberdade	Níveis de Significância, α em %			
	10	5	2,5	1
1	2,706	3,841	5,024	6,635
2	4,605	5,991	7,378	9,210
3	6,251	7,815	9,348	11,345
4	7,779	9,488	11,143	13,277
5	9,236	11,070	12,832	15,086
6	10,645	12,592	14,499	16,812
7	12,017	14,067	16,013	18,475
8	13,361	15,507	17,535	20,090
9	14,684	16,919	19,023	21,666
10	15,987	18,307	20,483	23,209
11	17,275	19,675	21,920	24,725
12	18,549	21,026	23,337	26,217
13	19,812	22,362	24,736	27,688
14	21,064	23,685	26,119	29,141
15	22,207	24,996	27,488	30,578
16	23,642	26,296	28,845	32,000
17	24,769	27,587	30,191	33,409
18	25,989	28,869	31,526	34,805
19	27,204	30,144	32,852	36,191
20	28,412	31,410	34,170	37,566

3.4.2. Teste de aderência por Kolmogorov-Smirnov

Este teste foi introduzido por Kolmogorov e Smirnov (1933), citados por ASSIS (1996), para avaliação de uma específica e bem conhecida distribuição $F(x)$, a dados provenientes de uma distribuição $S(x)$.

Sua vantagem sobre o teste de χ^2 é que ele pode ser aplicado, sem restrição, em amostras pequenas. Além disso, ele pode tratar os dados individualmente, não perdendo informações devido a agrupamentos, como no teste de χ^2 .

Como metodologia para sua aplicação, pode-se considerar $F(x)$ a proporção dos valores esperados menores ou iguais a x e $S(x)$ a proporção dos valores observados menores ou iguais a x , em que D_{obs} é módulo do desvio máximo observado:

$$D_{obs} = \text{Max} | F(x) - S(x) | \quad 55$$

Para testar em nível α de significância específico para, pelo menos, um valor de x , compara-se D_{obs} com D_{tab} (D_{tab} é o desvio máximo tabelado); se D_{obs} for menor, existe concordância entre as freqüências observadas e esperadas, a amostra provém de uma população que segue a distribuição de probabilidade sob teste.

Pode-se considerar os limites aproximados de $D_{tab} = \frac{k}{\sqrt{N_{Obs}}}$, em que os valores de k são dados pela tabela a seguir.

Tabela 3 - Valores críticos de D_{tab} para o teste de Kolmogorov-Smirnov

	Nível de Significância, α em %				
	20	10	5	2	1
K	1,07	1,14	1,22	1,37	1,63

3.5. Análise de agrupamento

A técnica de classificação denominada análise de agrupamento (ou análise de aglomerados ou *cluster analysis*), em geral, requer análise de componentes principais. A análise de agrupamento constitui uma metodologia numérica multivariada, com o objetivo de obter uma estrutura classificatória ou de reconhecimento de grupos, em que, dado um conjunto de “p” variáveis, no caso estações pluviométricas, as quais são diferenciadas segundo “n” características, no caso médias pluviométricas, pode-se obter um esquema que possibilite reunir as variáveis em um número de grupos tal que exista homogeneidade dentro e heterogeneidade entre os grupos (CRUZ e REGAZZI, 1997).

O processo de agrupamento envolve basicamente duas etapas. A primeira relaciona-se com a estimação de uma medida de dissimilaridade entre os indivíduos, e a segunda, com a adoção de uma técnica de agrupamento para a formação dos grupos.

3.5.1. Medidas de dissimilaridade

Um conceito fundamental na utilização das técnicas de análise de agrupamento é a escolha de um critério que meça a distância entre dois objetos, ou que quantifique o quanto eles são parecidos. Se os dados estão todos no mesmo padrão de medida, a variabilidade de cada variável é homogênea ou quase homogênea, podendo ser usados os dados originais. Os índices “i” e “j” representam as linhas e colunas da matrix de variáveis e características, respectivamente, e p o número de variáveis.

Entre as medidas de dissimilaridade podem ser citadas:

a) Distância euclidiana

A distância euclidiana entre as variáveis x_i e x_j é dada por

$$d_{ij} = \left[\sum_{j=1}^p (x_{ij} - x_{ij})^2 \right]^{\frac{1}{2}}$$

56

b) Distância euclidiana média

A distância euclidiana cresce à medida que cresce o número de variáveis. Uma maneira de contornar esse problema é dividir esse valor pela raiz quadrada do número de variáveis, obtendo-se a distância euclidiana média.

$$\Delta_{ij} = \frac{1}{\sqrt{p}} d_{ij} \quad 57$$

c) Distância de Mahalanobis

A distância de Mahalanobis, entre as variáveis x_i e x_j , é dada por

$$D_{ij} = \sum_{j=1}^p (x_{ij} - x_{i'j})^2 \quad 58$$

em que x_{ij} representa o valor da i -ésima variável referente à j -ésima característica e p o número de variáveis.

3.5.2. Métodos de agrupamento

De acordo com EVERITT (1980), o método de agrupamento pode ser classificado em cinco grupos, dos quais os mais importantes são:

1. Método hierárquico: os indivíduos são reunidos em grupos, e o processo repete-se em diferentes níveis, até formar um diagrama em árvore (dendograma).
2. Método de otimização: os grupos são formados pela otimização de um critério de agrupamento. Os grupos são mutuamente exclusivos, formando uma partição do conjunto de entidades.

Entre os métodos hierárquicos, os processos aglomerativos podem ser:

1. Método do vizinho mais próximo ou método do encadeamento simples ou, ainda, "simple linkage", em que o coeficiente de distância entre duas variáveis x_i e x_j será dado por

$$d_{ij} = \min d_{ij} \quad 59$$

2. Método do vizinho mais distante ou método do encadeamento completo ou, ainda, “complete linkage”, em que o coeficiente de distância entre duas variáveis x_i e x_j será dado por

$$d_{ij} = \max d_{ij}$$

60

Entre os métodos de otimização tem-se o de Tocher, citado por RAO (1952), em que se adota o critério de que a média das medidas de dissimilaridade dentro de cada grupo deve ser menor que as distâncias médias entre quaisquer grupos. O método requer a obtenção da matriz de dissimilaridade, sobre a qual é identificado o par de indivíduos mais similar; estes indivíduos formarão o grupo inicial. A partir daí é avaliada a possibilidade de inclusão de novos indivíduos, adotando-se o critério anteriormente citado.

$$\frac{d(\text{grupo})_i}{g} \leq \theta, \text{ inclui a variável } x_i \text{ no grupo atual.} \quad 61$$

$$\frac{d(\text{grupo})_i}{g} > \theta, \text{ a variável } x_i \text{ não deve ser incluída no grupo atual.} \quad 62$$

em que g é o número de elementos que constituem o grupo que está sendo considerado e θ é a maior distância entre as menores.

4. MATERIAL E MÉTODOS

Para realização do presente trabalho, foram utilizados mapas digitalizados do Estado de Minas Gerais, situado na Região Sudeste do Brasil, localiza-se entre os paralelos de 14°13'57" e 22°55'22" de latitude Sul e os meridianos de 39°51'23" e 51°02'45" a Oeste de Greenwich, na escala 1:1.500.000, disponíveis no *site* do Geominas (www.geominas.org.br). Os dados de precipitação foram fornecidos pela Agência Nacional de Águas e Energia Elétrica (ANEEL) e pelo Instituto Nacional de Meteorologia (INMET).

Para edição dos mapas e representação dos dados foram utilizados o *software* ArcView®, versão 3.1, que é um Sistema de Informações Geográficas desenvolvido pelo *Environmental Systems Research Institute*. Para cálculo dos totais de precipitação e aplicação das distribuições de probabilidade, foram desenvolvidos programas escritos em Object Pascal para Delphi®, versão 5.0, desenvolvido pela *Inprise Corporation*.

As etapas do trabalho estão descritas a seguir:

4.1. Seleção das estações

O banco de dados pluviométricos atualmente disponível para Minas Gerais contém informações de precipitação diária de aproximadamente 943 estações da ANEEL e 39 do INMET.

A *Food and Agricultural Organization* (FAO) preconiza que o número mínimo de anos de dados climáticos para análise é de 30 anos. A *World Meteorological Organization* (WMO), em sua nota técnica nº 81, recomenda o *run test*, como definido no item 3.2. Entretanto, foi considerado que a não-aprovação de uma estação no teste não implicaria na sua retirada das análises, pois serviria apenas como referência da situação dos dados.

Foi desenvolvido um programa escrito em Object Pascal para Delphi, para aplicação do teste, análise gráfica, correção de datas, limites inferior e superior de dados em todas as estações, a fim de selecionar as que estavam em condições de análise. Para o processo de interpolação de dados foram utilizadas estações vizinhas ao estado, a fim de evitar tendenciosidade nesse processo nos limites do estado. Cada estação possui as seguintes informações para representá-la (Tabela 4):

Tabela 4 - Informações associadas a cada estação

Campo	Descrição
Código	código da estação segundo o órgão responsável
Nome	nome da estação segundo o órgão responsável
Município	município mais próximo da estação ou a que pertence
Região	região do estado a estação está localizada
Latitude	latitude da estação em graus decimais
Longitude	longitude da estação em graus decimais
Período	ano em que se iniciou e encerrou a coleta dos dados
Anos de dados	número de anos de dados que a estação possui

4.2. Totais de precipitação

Foram considerados, para fins de análise, o total médio e o número médio de dias com chuva, para os períodos anual, chuvoso, seco e mensal de precipitação. Como período chuvoso consideraram-se os meses de janeiro, fevereiro, março, outubro, novembro e dezembro, e, como período seco, os

meses de abril, maio, junho, julho, agosto e setembro. Foram considerados dias com chuva aqueles que apresentaram precipitação maior ou igual a 0,1 mm.

Foi elaborado um banco de dados com as informações da Tabela 5, em que cada registro representava uma estação analisada. Nesses cálculos os anos bissextos foram desconsiderados.

Foi utilizado o editor *Map Calculator* do ArcView 3.1, módulo *Spatial Analyst*, para seleção dos campos da Tabela 5, utilizados na geoespacialização desses totais de precipitação.

Tabela 5 - Valores calculados de cada estação para análise dos totais de precipitação

Campo	Descrição
TMA anual	Total médio de precipitação anual
TMPChuvoso	Total médio de precipitação do período chuvoso
TMPSeco	Total médio de precipitação do período seco
TMMensalMes1	Total médio de precipitação mensal de janeiro
...	
TMMensalMes12	Total médio de precipitação mensal de dezembro
TMDiasAnoCC	Total médio de dias do ano com chuva
TMDiasPChuvosoCC	Total médio de dias do período chuvoso com chuva
TMDiasPSecoCC	Total médio de dias do período seco com chuva
TMDiasCCMes1	Total médio de dias com chuva do mês de janeiro
...	
TMDiasCCMes12	Total médio de dias com chuva do mês de dezembro

Na interpolação do dados, utilizou-se um interpolador linear do *ArcView*, IDW, em que o peso da célula a ser interpolada é dado por uma média ponderada que utiliza o peso dos pontos de controles mais próximos (definido pelo usuário), ponderados pelo inverso da distância elevado a um expoente “m” (definido pelo usuário), ou seja:

$$G_i = \frac{\sum_{i=1}^n \left(\frac{x_i}{d_i^m} \right)}{\sum_{i=1}^n \left(\frac{1}{d_i^m} \right)} \quad 65$$

em que,

- G_i = peso estimado no ponto interpolado;
- x_i = peso do i -ésimo ponto amostrado;
- d_i = distância euclidiana do i -ésimo ponto amostrado ao ponto atual
- m = expoente da função de distância euclidiana.

Foram considerados os 12 pontos mais próximos para interpolação e a segunda potência empregada no cálculo da distância euclidiana, que são índices indicados por outros trabalhos na área (ASSAD, 1994). Foram utilizadas estações vizinhas ao Estado para evitar tendenciosidade no processo de geoespacialização.

Os mapas vetoriais de Minas Gerais, utilizados no ArcView, foram obtidos por meio de *download* disponível no *site* Geominas, sob responsabilidade da Secretaria de Estado da Casa Civil do Governo de Minas Gerais. Esses mapas foram elaborados pela Prodemge (Companhia de Processamento de Dados do Estado de Minas Gerais), a partir do mapa geopolítico de Minas Gerais, escala 1:1.500.000.

Os mapas foram utilizados para avaliar características espaciais e temporais dos totais de precipitação e do número médio de dias em cada período.

4.3. Distribuições de probabilidade

Foram considerados, para fins de aplicação das distribuições, os valores descritos na Tabela 6. Consideraram-se somente dias com chuva, ou seja, aqueles com precipitação maior ou igual a 0,1 mm; dias com precipitação igual a zero foram desconsiderados, devido ao uso de função logarítmica para cálculo dos parâmetros de algumas distribuições. Os anos bissextos foram desconsiderados.

Tabela 6 - Valores calculados de cada estação para aplicação das distribuições de probabilidade

Campo	Descrição
D1DecMes1 a 12	Precipitação diária 1 ^o decêndio do mês de janeiro a dezembro
D2DecMes1 a 12	Precipitação diária 2 ^o decêndio do mês de janeiro a dezembro
D3DecMes1 a 12	Precipitação diária 3 ^o decêndio do mês de janeiro a dezembro
DMensalMes1 a 12	Precipitação diária mensal do mês de janeiro a dezembro
T1DecMes1 a 12	Total de precipitação 1 ^o decêndio de janeiro a dezembro
T2DecMes1 a 12	Total de precipitação 2 ^o decêndio de janeiro a dezembro
T3DecMes1 a 12	Total de precipitação 3 ^o decêndio de janeiro a dezembro
TMensalMes1 a 12	Total de precipitação do mês de janeiro a dezembro

Anteriormente, foram citados alguns métodos para cálculo das distribuições e dos seus parâmetros; estes métodos foram testados para escolha dos que apresentarem melhor ajuste. As equações utilizadas, em cada um dos testes, são apresentadas a seguir.

4.3.1. Distribuição exponencial

Utilizou-se a função cumulativa de probabilidade apresentada pela Equação 2.

4.3.2. Distribuição gama

O cálculo da distribuição gama se compõe de três partes.

Cálculo dos parâmetros da distribuição, o que é feito pelo método da máxima verossimilhança, item 3.3.2.1, ou método de Greenwood e Durand (1960), item 3.3.2.2.

Para a resolução aproximada da função gama a fim de selecionar a equação que melhor estima os valores da função gama, foram construídos algoritmos com as equações citadas no item 3.3.2.3. Em seguida, foram

selecionados intervalos para teste das equações em todo o campo de abrangência da função. Para cálculo da função gama foram utilizados um *software* matemático e um algébrico, que apresentam precisão de 15 dígitos, sendo considerados de referência.

Feito o cálculo da integral definida da função cumulativa de probabilidade da variável aleatória, Equação 4, pelas técnicas de integração numérica, como a de Gauss, e a utilização de uma expansão em série apresentada pela Equação 32.

4.3.3. Distribuição log-normal

Utilizou-se a distribuição de densidade log-normal a dois e três parâmetros, conforme a Equação 35, definida no item 3.3.3

4.3.4. Distribuição normal

Foram feitos testes para o processo de obtenção da probabilidade pelo cálculo da integral, utilizando o método de Gauss, e aquele obtido por uma função aproximadora. Pela função original tem-se a Equação 37 e a função aproximadora a Equação 39.

4.3.5. Distribuição Weibull

A função de densidade acumulada usada foi a da Equação 41 definida no item 3.3.5. É normalmente chamada de Weibull a dois parâmetros, pois o parâmetro α assume um valor mínimo na série de dados a ser aplicada. Para o cálculo dos parâmetros β e γ , foi usado o método da máxima verossimilhança, Equações 50 e 51 pelas razões já expostas naquele item.

4.4. Testes de aderência

Para verificar o nível de aderência entre a probabilidade observada e a estimada, foram utilizados os testes de qui-quadrado e Kolmogorov-Smirnov, conforme apresentados nos itens 3.4.1 e 3.4.2. Em que foram utilizados os níveis de 5% e 20% de significância, respectivamente.

4.5. Análise de agrupamento

Foram utilizadas 624 estações nesta análise. Foi desenvolvido um programa em Object Pascal para Delphi com a teoria apresentada no item 3.5. Foi utilizado um *software* gráfico, para construção do dendograma, com apenas 212 estações, que apresentam dados acima de 30 anos. Foram utilizadas as seguintes informações constantes da Tabela 7.

Tabela 7 - Informações para análise de agrupamento

Campo	Descrição
TMA anual	Total médio anual
TMPChuvoso	Total médio do período chuvoso
TMPSeco	Total médio do período seco
TMMensalMes1	Total médio mensal do mês de janeiro
...	...
TMMensalMes12	Total médio mensal do mês de dezembro

Foi realizada uma combinação entre as medidas de dissimilaridade e os processos de agrupamento, o que apresentasse melhor distribuição de grupos, dentre os avaliados, seria escolhido para aplicação.

Partindo desse princípio, procurou-se avaliar a existência de grupos homogêneos de precipitação média para todo o Estado.

O dendograma obtido foi analisado, a fim de escolher um índice de salto adequado para que o número de grupos não ficasse muito grande, dificultando a análise, ou muito pequeno, encobrindo características de algumas estações, já que essa inferência é normalmente visual (REGAZZI, 1999).

5. RESULTADO E DISCUSSÃO

5.1. Seleção das estações

Dentre as 982 estações disponíveis selecionou-se, para estudos, 603 estações da ANEEL e 21 do INMET. Para o processo de interpolação nos limites do Estado, foram utilizadas 190 estações da ANEEL.

Inicialmente, foram consideradas somente estações com mais de 30 anos de dados diários, conforme orienta a FAO. Entretanto, tendo em vista que a Tabela 8 apresenta cerca de 545 estações que possuem entre 10 e 29 anos de dados diários, o que representa cerca de 67% do total.

Julgou-se então importante para o processo de espacialização a utilização destas, pois, caso contrário, haveria regiões do Estado sem nenhuma fonte de dados que as representasse. Por outro lado, na análise das distribuições de probabilidade foram consideradas somente as estações com 30 anos de dados ou mais.

A Tabela 1A apresenta a identificação das 814 estações utilizadas na análise. A Figura 1 ilustra a disposição geográfica das estações aptas para análise.

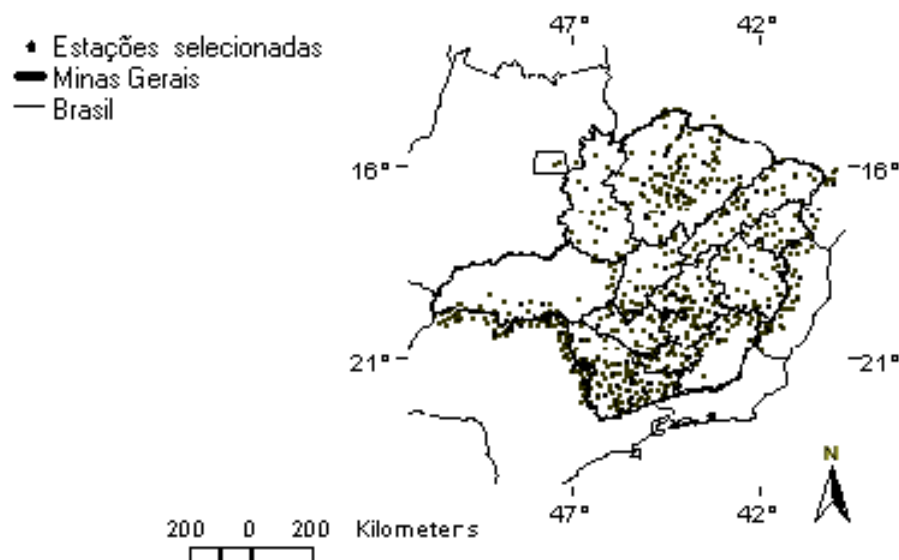


Figura 1 - Localização das estações estudadas no Estado de Minas Gerais.

Pode-se observar, na Figura 1, que ao norte da região do Triângulo Mineiro e ao sul da Zona da Mata, o número de estações é limitado.

Ao sul da Bahia e leste de Goiás, limites com o Estado de Minas Gerais, segundo os dados da Tabela 8, há cerca de 19 e 3 estações em condições de serem analisadas, respectivamente; o que representa um número muito reduzido de estações com poucos anos de dados em uma região de potencial agrícola. Isto implica em resultados de pouca precisão quando da aplicação dos processos de interpolação e análise.

A FAO preconiza que o número mínimo de anos de dados diários para uma análise de dados meteorológicos é de 30 anos. Entretanto, pode-se notar que, na Tabela 8, somente 33% das estações atendem a estas condições, e essas informações, devido à sua própria natureza, não podem ser recuperadas. Outro aspecto importante a ser observado é a descontinuidade da obtenção dos dados. Verificou-se que 28% das estações não apresentavam dados contínuos ao longo do seu período de obtenção, havendo falha de 1 até 45 anos entre informações de uma mesma estação. Considerando o período entre 1960 e 1990, somente 14% das estações teriam dados de forma contínua, o que limita o levantamento das relações de eventos entre as

estações, pois comportamentos importantes poderiam ser correlacionados ao longo dos anos, caso os dados tivessem, em sua totalidade, maior continuidade e a rede meteorológica, melhor distribuição.

Tabela 8 Regiões estudadas e total de dados disponíveis

Região	Nº de estações	Nº de estações com período entre 10 e 29 anos	Nº de estações com período entre 30 e 74 anos
Vale do Mucuri	14	4	10
Noroeste de MG	25	22	3
Triângulo	26	23	3
Central Mineira	26	20	6
Campo das Vertentes	30	12	18
Oeste de MG	32	14	18
Zona da mata	32	19	13
Vale do Rio Doce	36	22	14
Jequitinhonha	44	31	13
Sul de MG	112	58	54
Metropolitana BH	114	72	42
Norte de MG	133	115	18
Limites de MG			
Goiás	3	3	0
Bahia	19	14	5
Espírito Santo	40	26	14
São Paulo	128	90	38
Total	814	545	269

5.2. Totais médios de precipitação

Foram utilizados três tipos de legendas na espacialização, cada uma com 10 classes de precipitação: a) com amplitude de 50 mm, usada na precipitação média do período seco e dos meses com chuva; b) com amplitude de 15 mm, usada na precipitação dos meses secos; e c) com amplitude de 200 mm, usada na precipitação anual e média do período com chuva. As cores das legendas foram escolhidas com o objetivo de proporcionar melhor distinção entre os intervalos, não havendo correlação com a intensidade do fenômeno que representam.

A Figura 2a (precipitação média anual) apresenta três classes de precipitação bem definidas e distribuídas ao longo do Estado. Na extremidade norte do estado, região mais seca, tem-se uma faixa com precipitação entre 600 e 800 mm. Até a região central do Estado, tem-se uma classe de 800 a 1.000 mm. Em seguida, 1.000 a 1.200 mm, iniciando na região noroeste e terminando ao sul da Zona da Mata. Finalmente no sul do estado, tem-se uma grande área com classe de 1.200 a 1.400 mm, que se estende até o norte de São Paulo. Nesta região há pequenas zonas onde a precipitação alcança valores entre 1.400 a 1.600 mm.

A Figura 2b (precipitação média do período com chuva) apresenta aproximadamente a mesma distribuição da média anual, contendo três classes de precipitação bem definidas e distribuídas ao longo do estado, o que mostra que o período com chuva que compreende os meses de janeiro, fevereiro, março, outubro, novembro e dezembro contribui consideravelmente para a formação desses totais de precipitação anual, deixando para o período seco uma quantidade bem pequena de contribuição: a) ao norte do estado, tem-se de 600 a 800 mm; b) uma faixa que vai da região Noroeste até o sul da Zona da Mata, com 800 a 1.200 mm; e c) no sul do Estado uma área com 1.000 a 1.200 mm, com pequenas zonas que alcançam 1.200 a 1.400 mm. De forma geral, pode-se notar que a classe média de precipitação, tanto anual quanto no período com chuva, fica em torno de 800 a 1.400 mm em cerca de 85% do estado.

Na Figura 2c (a precipitação média do período seco) que compreende os meses de abril, maio, junho, julho, agosto e setembro, caracteriza-se por possuir de maneira bem irregular três classes de precipitação; ao norte do estado, compreendendo as regiões Noroeste, Norte e Jequitinhonha, tem-se de 50 a 150 mm; numa faixa estreita no meio do estado uma região com 150 a 200 mm; ao sul do Estado, compreendendo somente as regiões sul do Triângulo, Sudoeste, Oeste e Campo das Vertentes, uma classe de 200 a 250 mm; o período seco apresenta totais de precipitação bem inferiores aos totais anuais e do período com chuva, caracterizando assim uma natural divisão do estado em dois períodos bem distintos, um com chuvas médias próximas de 1.200 mm e outro com 200 mm. Dentro do período seco, bem ao norte do vale do Jequitinhonha e Vale do Mucuri, tem-se uma classe com 200 a 250 mm dentro de uma região com totais médios

de 100 a 150 mm; isto se deve possivelmente à influência da mata atlântica, que proporciona estes totais nesta região.

Ao analisar as relações entre as precipitações mensais do mês de janeiro e dezembro (Figuras 3a e 4f), pode-se notar uma semelhança na distribuição média mensal desses meses. Para os meses de janeiro e dezembro, tem-se aproximadamente 80% do estado com precipitação entre 200 e 300 mm, ficando somente a extremidade nordeste entre 100 e 200 mm.

Ao observar os meses do período seco, pode-se notar um comportamento de distribuição de precipitação muito semelhante nos meses de junho, julho e agosto (Figuras 3f, 4a e 4b), que são os meses mais carentes de precipitação em todo o estado, em que as classes são de 0 a 15 mm, na direção centro-norte. No mês de abril há diminuição das classes de chuva do mês de março, e, no mês de setembro, aumento para as classes de outubro. Nota-se que nos meses com chuva a precipitação está em torno de 200 mm, e no período seco ela não ultrapassa 100 mm.

De forma geral (Figuras 3a a 4f), pode-se notar que no espaço, caminhando-se do norte ao sul do estado, as precipitações médias aumentam nesse sentido. No tempo, de janeiro a dezembro, estas têm a forma de "U", caracterizando grandes precipitações nos três meses iniciais e finais do ano e precipitações mínimas nos demais meses do ano.

5.3. Número médio de dias com chuva

Foram utilizados dois tipos de legendas, cada uma com 10 classes de número médio de dias com chuva: uma com amplitude de três dias e outra com amplitude de 18 dias. Foi usado, na legenda, o mesmo padrão de cores, dos totais de precipitação, sem relação com o número médio de dias com chuva.

O número médio de dias com chuva é um índice importante a ser obtido de um banco de informações pluviométricas, pois ele pode determinar o nível de utilização de equipamentos agrícolas, bem como o manejo de outras atividades agrícolas.

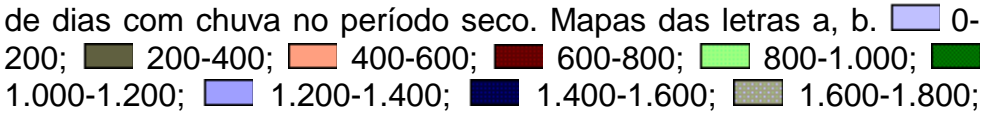





























Na Figura 2d (número médio de dias com chuva no ano) pode-se notar que, na região norte, há quatro classes de número médio de dias, Na Figura 2e podem-se notar perfeitamente três classes bem distintas: ao norte, de 36 a 54 dias; ao centro, de 54 a 72 dias; e ao sul, maior parte, de 72 a 90 dias. Na Figura 2f (número médio de dias com chuva no período seco) há somente duas classes, do norte até a região central do Estado, uma classe de 0 a 18 dias; e, do centro para o sul, outra classe, de 18 a 36 dias.

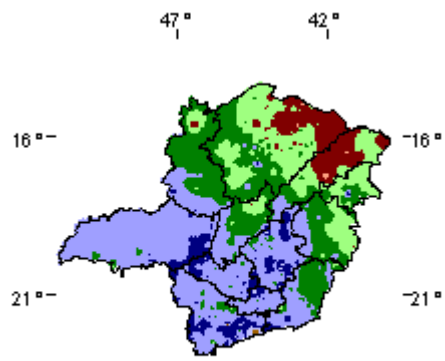
Foram utilizadas 814 estações para espacialização desses valores; mesmo com essa densidade, verifica-se que há separação significativa entre um valor e outro de número médio de dias, caracterizando um comportamento homogêneo na região de ocorrência desses valores.

Para o período com chuva, tem-se uma intensidade média de chuvas de cerca de 1.200 mm, distribuídos em aproximadamente 70 dias; no período seco são 200 mm para 18 dias aproximados, o que indica uma concentração de chuva em poucos dias, pois cada período tem aproximadamente 180 dias. Isto pode exigir algumas técnicas de irrigação ou drenagem, dependendo do período do ano e da cultura, o que aumenta os custos da produção.

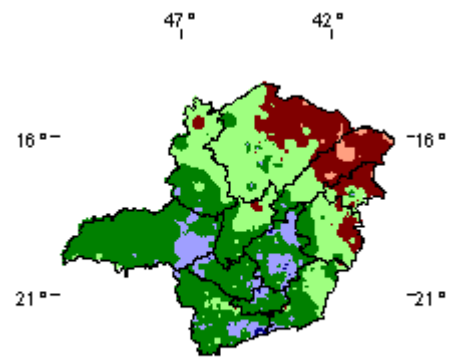
Observando as Figuras 5a a 6f (número médio de dias com chuva nos meses de janeiro a dezembro) pode-se notar que os três meses iniciais e finais do ano apresentam número de dias maiores, diminuindo nos demais meses do ano, formando uma curva em forma de "U". Tanto na precipitação média mensal quanto no número médio de dias com chuva, este, comportamento, distribuição média em forma de "U" para os meses do ano, é bem definido. O conjunto formado entre os meses de fevereiro e março; maio e setembro; e junho, julho e agosto possui aproximadamente uma mesma distribuição de número médio de dias com chuva, havendo alteração em algumas extremidades do estado. Comparando os mapas de totais médios de precipitação para estes mesmos meses, tem-se pequena variação na precipitação para um mesmo número médio de dias.

O relacionamento entre os mapas de médias de precipitação e número médio de dias com chuva nos respectivos meses do ano proporcionará ao administrador rural uma melhor visão da distribuição destas duas variáveis, podendo fazer as considerações em seu projeto de maneira segura.

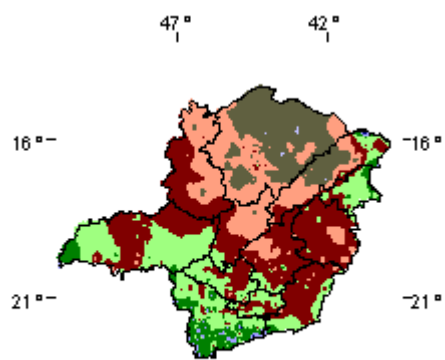
Figura 2 - a) precipitação média anual em mm; b) precipitação média do período com chuva em mm; c) precipitação média do período seco em mm; d) número médio de dias com chuva no ano; e) número médio de dias com chuva no período com chuva; f) número médio de dias com chuva no período seco. Mapas das letras a, b.  0-200;  200-400;  400-600;  600-800;  800-1.000;  1.000-1.200;  1.200-1.400;  1.400-1.600;  1.600-1.800;  1.800-2.050. Mapa c:  0-50;  50-100;  100-150;  150-200;  200-250;  250-300;  300-350;  350-400;  400-450;  450-500. Mapas d, e, f.  0-18;  18-36;  36-54;  54-72;  72-90;  90-108;  108-126;  126-144;  144-162;  162-180.



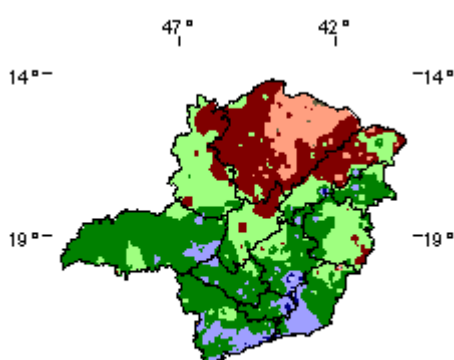
a



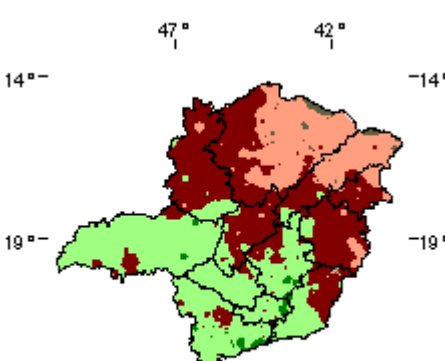
b



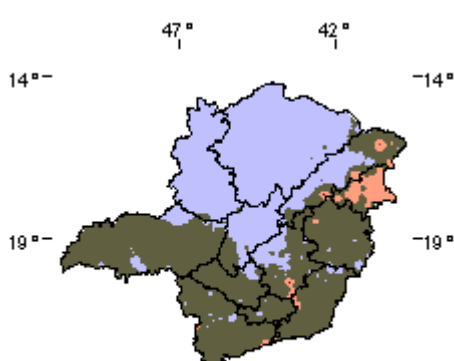
c























d



e



f

Figura 3 - a) precipitação média do mês de janeiro em mm; b) precipitação média do mês de Fevereiro em mm; c) precipitação média do mês de março em mm; d) precipitação média do mês de abril em mm; e) Precipitação média do mês de maio em mm; f) precipitação média do mês de junho em mm. Mapas das letras a, b, c:  0-50;  50-100;  100-150;  150-200;  200-250;  250-300;  300-350;  350-400;  400-450;  450-500. Mapas d, e, f:  0-15;  15-30;  30-45;  45-60;  60-75;  75-90;  90-105;  105-120;  120-135;  135-150

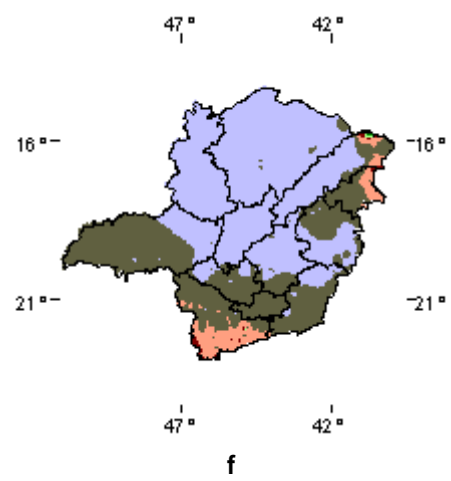
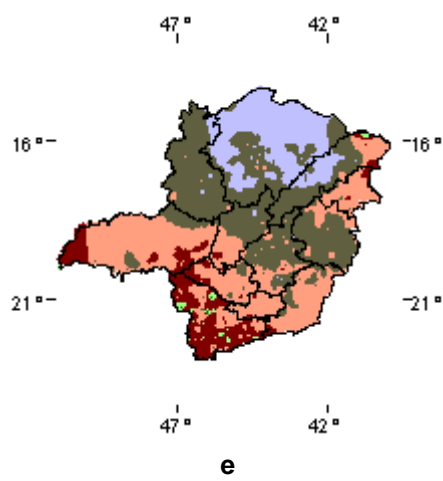
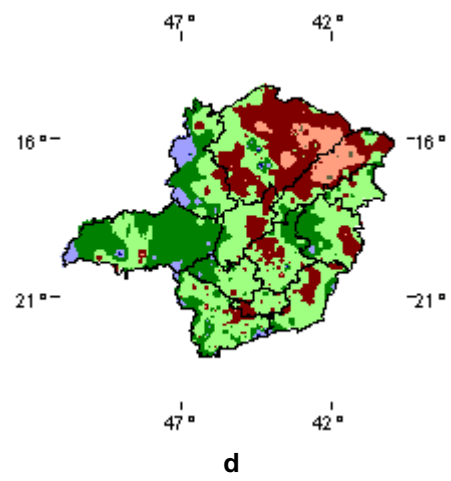
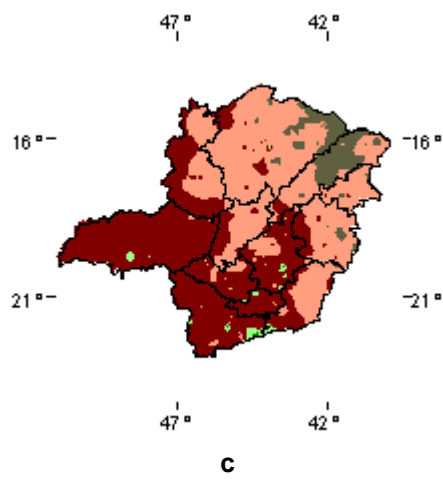
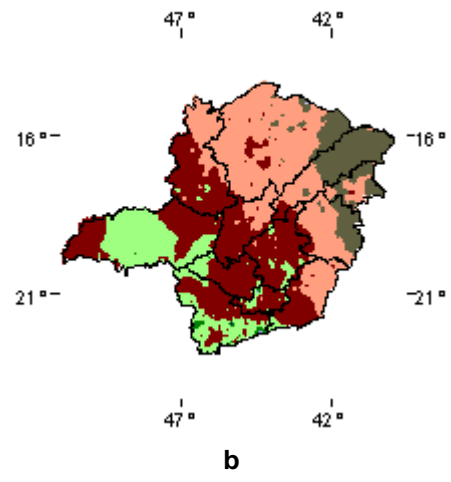
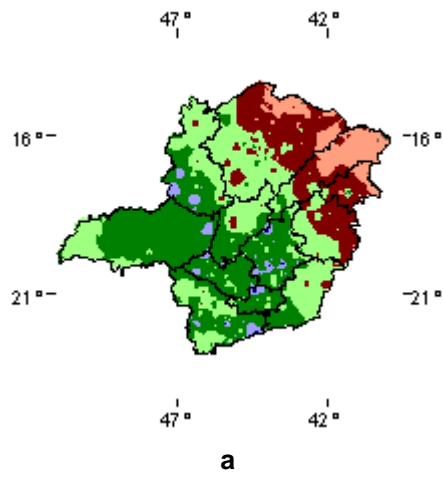
































Figura 4 - a) precipitação média do mês de julho em mm; b) precipitação /9 média do mês de agosto em mm; c) precipitação média do mês de setembro em mm; d) precipitação média do mês de outubro em mm; e) precipitação média do mês de novembro em mm; f) precipitação média do mês de dezembro em mm. Mapas das letras a, b, c:  0-15;  15-30;  30-45;  45-60;  60-75;  75-90;  90-105;  105-120;  120-135;  135-150. Mapas d, e, f.  0-50;  50-100;  100-150;  150-200;  200-250;  250-300;  300-350;  350-400;  400-450;  450-500. Mapas d, e, f:  0-15;  15-30;  30-45;  45-60;  60-75;  75-90;  90-105;  105-120;  120-135;  135-150.

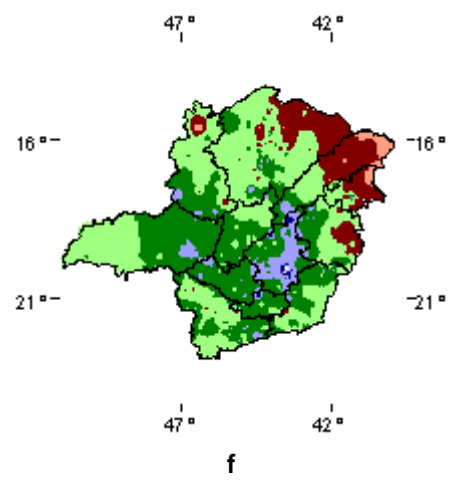
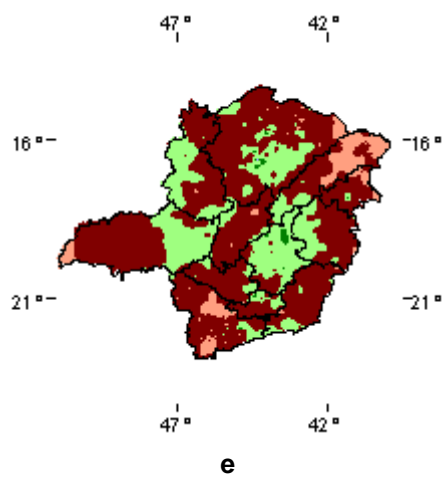
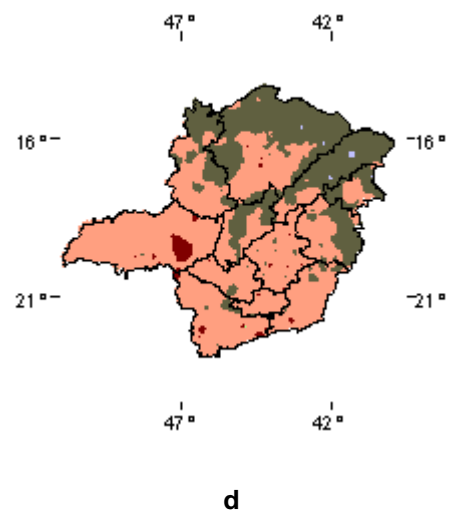
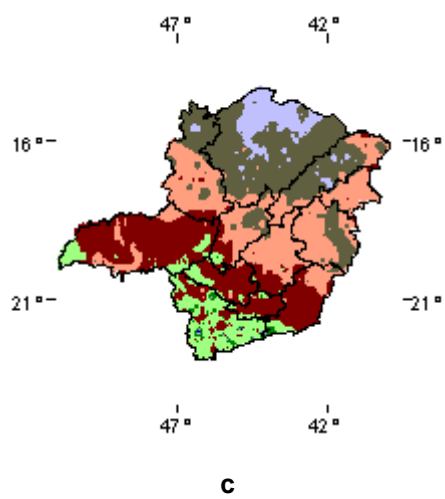
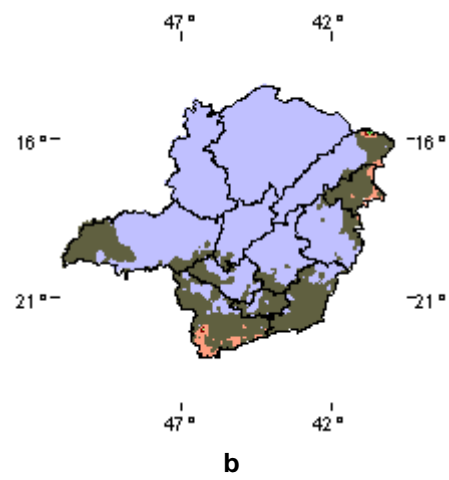
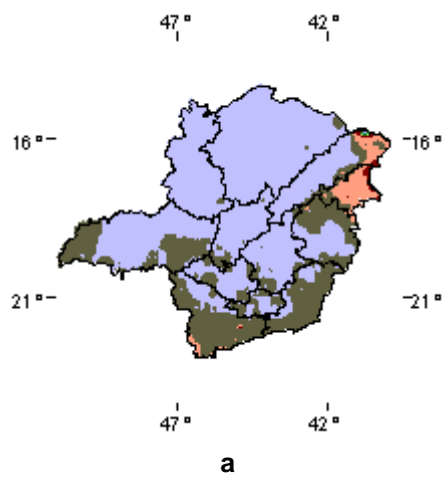
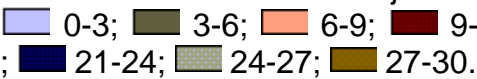





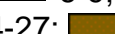
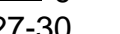

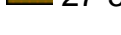
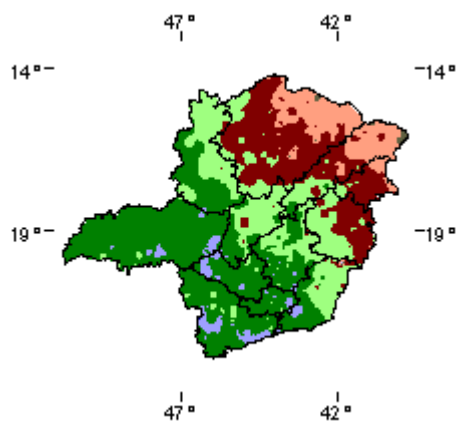
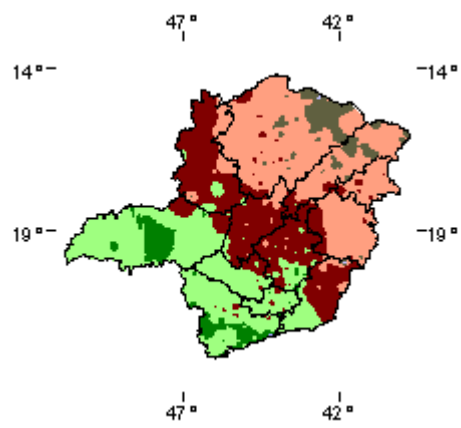


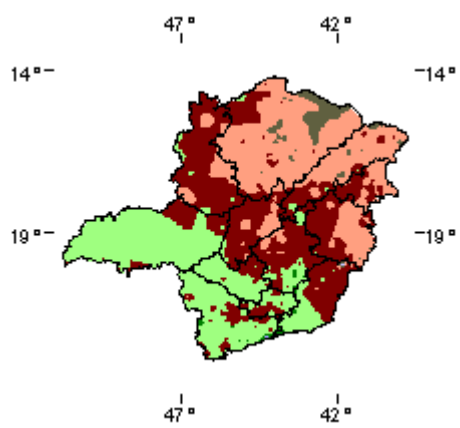
Figura 5 - a) número médio de dias com chuva no mês de janeiro; b) número médio de dias com chuva no mês de fevereiro; c) número médio de dias com chuva no mês de março; d) número médio de dias com chuva no mês de abril; e) número médio de dias com chuva no mês de maio; f) número médio de dias com chuva no mês de junho. Mapas das letras a, b, c, d, e, f:  0-3;  3-6;  6-9;  9-12;  12-15;  15-18;  18-21;  21-24;  24-27;  27-30.



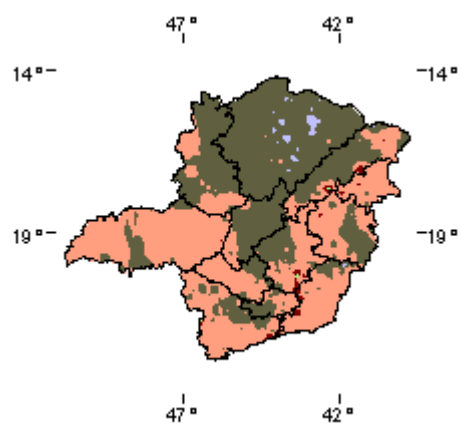
a



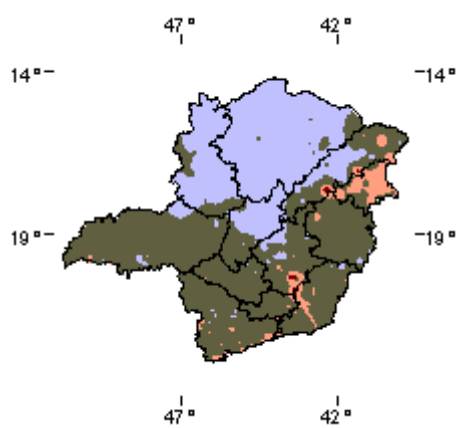
b



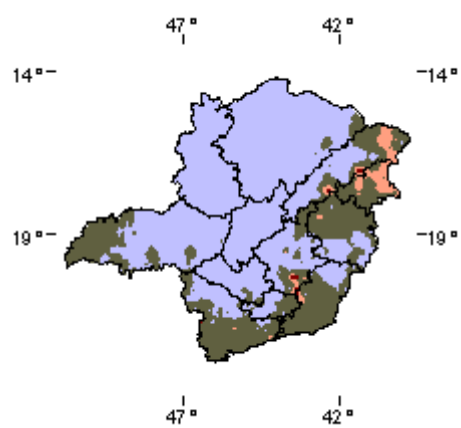
c



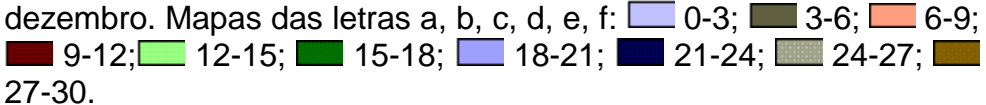









d

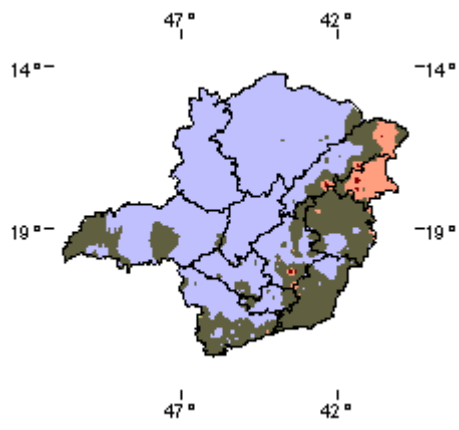


e

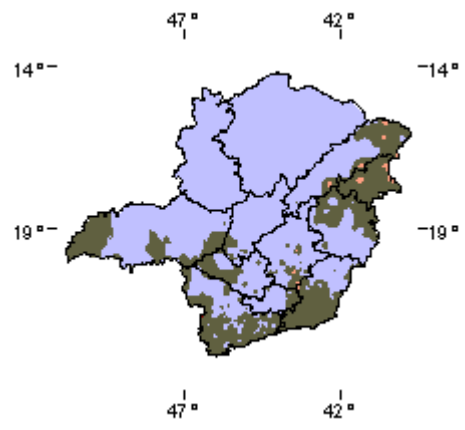


f

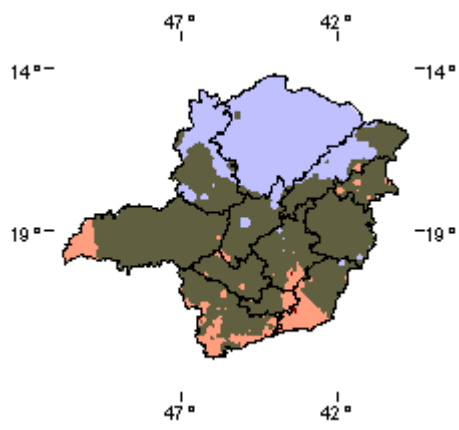
Figura 6 - a) número médio de dias com chuva no mês de julho; b) número médio de dias com chuva no mês de agosto; c) número médio de dias com chuva no mês de setembro; d) número médio de dias com chuva no mês de outubro; e) número médio de dias com chuva no mês de novembro; f) número médio de dias com chuva no mês de dezembro. Mapas das letras a, b, c, d, e, f:  0-3;  3-6;  6-9;  9-12;  12-15;  15-18;  18-21;  21-24;  24-27;  27-30.



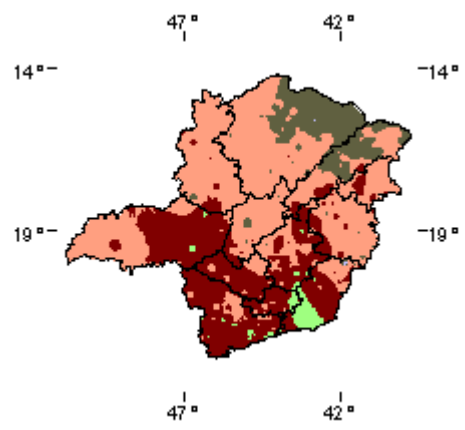
a



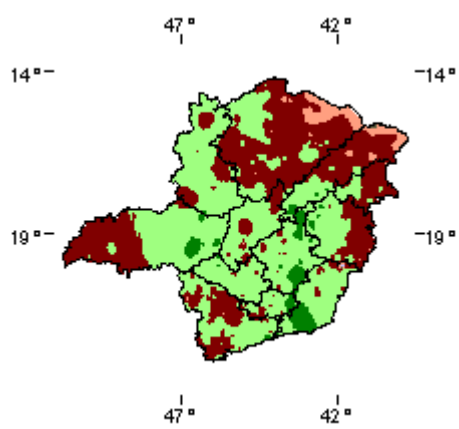
b



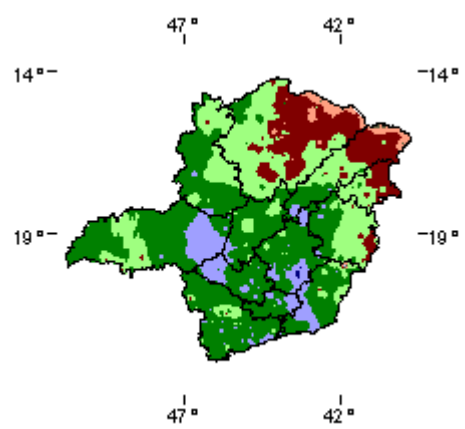
c



d



e



f

5.4. Distribuições de Probabilidades

Nos tópicos seguintes são apresentadas as análises de algumas distribuições de densidade, procurando destacar algumas características importantes, independentemente de elas terem tido ou não bom desempenho na análise geral.

5.4.1. Distribuição exponencial

A distribuição de freqüência de dados pluviométricos apresenta uma assimetria em forma de "J" invertido, típica da distribuição exponencial.

Os dados se concentram nas classes iniciais e diminuem a concentração nas finais. Pode-se questionar o fato de que, mesmo tendo naturalmente os dados, esta distribuição de freqüência (distribuição exponencial) não é a que melhor estima os valores observados para alguns períodos. Isto se deve à capacidade do modelo matemático em estimar cada classe de freqüência individualmente, quando ocorrem picos intermediários de freqüência.

Ao observar os dados das Tabelas 11 a 18, para a distribuição exponencial, verifica-se boa aderência nos meses com chuvas e diminuição nos meses secos, para os períodos decendiais totais e diários. Por outro lado, no período mensal total há uma inversão, tendo a melhor aderência os meses secos.

Considerando o erro relativo entre as probabilidades observadas e as estimadas, pode-se observar que as probabilidades estimadas no período com chuva são em grande parte superestimadas. A característica da distribuição exponencial é a de, partindo de um valor da classe inicial, ir estimando outros valores proporcionalmente menores, formando uma curva em forma de "J" invertido. Ao analisar a distribuição de classes de freqüência de algumas estações isoladamente, nota-se que as classes intermediárias têm valores maiores que as classes iniciais ou bem próximos desta, não formando uma curva em "J" invertido, mas sim algo próximo a um "M". A capacidade da

distribuição exponencial é limitada em estimar estes "vales" intermediários, superestimando-os, pois seu único parâmetro é a média; ao se efetuarem os testes de aderência, estas classes são reprovadas. Observa-se também que o desvio-padrão entre os três decêndios totais é pequeno, se comparado com o desvio-padrão do período mensal total. Com as classes de frequências das estações no período seco, esses picos intermediários não ocorrem, tendo estas a forma de um "J" invertido suave, o que reduz o erro relativo na estimação, aumentando o número de aderências.

5.4.2. Distribuição gama

Foram estimados os parâmetros da distribuição gama em aproximadamente 213 estações pluviométricas, que apresentam tipos de dados bem variados. Todos os cálculos dos parâmetros alfa e beta em que se utilizaram o método da máxima verossimilhança e método de Greenwood e Durand (1960) forneceram valores bem próximos, com erros na 10^a casa decimal, o que é uma precisão muito boa para as aplicações estatísticas e meteorológicas mais gerais.

A utilização dos dois métodos é igual para o cálculo do parâmetro beta, diferenciando apenas no processo de obtenção do parâmetro alfa. No método de Greenwood e Durand (1960) exige-se a verificação da faixa de abrangência da constante para aquele método, o que, em termos computacionais, pode tornar lento o cálculo para um número muito grande de estações, como é o caso do presente trabalho. Por outro lado, este método tem um limite superior, que não foi ultrapassado em nenhuma das estações, mas, estatisticamente, não se pode dizer que ele não será superado. Sem nenhuma restrição, o método da máxima verossimilhança apresenta-se mais adequado a todas as situações e foi o escolhido para cálculo dos parâmetros neste trabalho.

Para o teste da equação que apresenta a melhor *performance* no cálculo da função gama, foram feitos testes com diversos intervalos compreendidos entre 0 e 100 valores da variável alfa, com incremento de 0,0001. Os intervalos que apresentaram maior divergência entre as equações

foram os compreendidos entre 0 e 1, sendo os demais precisos até o 10^o dígito decimal. A Figura 7 apresenta, no eixo das ordenadas, os valores da função gama para alfa variando de 0,3 a 0,8. Foram desconsiderados os valores de 0,0001 a 0,2999, pois o erro entre as equações era muito grande e encobria os demais valores. Foi utilizada a legenda em que a assintótica refere-se à Equação 25; Stirling, à Equação 24; Assis P. A., à Equação 26; Polinômio, à Equação 21; e Software representa a referência, calculada pelo *software* matemático empregado na análise.

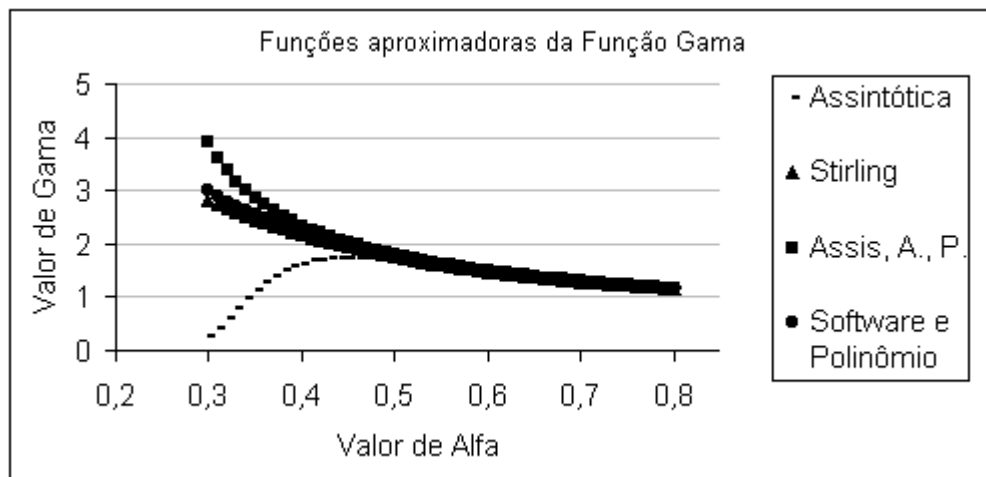


Figura 7 - Valores da função gama para alfa variando de 0,3 a 0,8, considerando diferentes funções aproximadoras.

Pode-se notar claramente que, usando a Equação 21, os valores calculados se aproximam muito do método tido como referência; enquanto o erro desta permanece próximo de zero, em alguns outros intervalos as demais equações superestimam ou subestimam os valores da função gama. Conclui-se que a utilização da Equação 21 é a melhor escolha para cálculo da função gama.

Como foi citado anteriormente, há dois métodos para cálculo da função cumulativa de probabilidade da distribuição gama: a) pela resolução da

equação de definição (Equação 4) utilizando método de integração de Gauss; ou b) usando uma função aproximadora (Equação 32).

Os dois métodos, em todos os cálculos, apresentaram igual precisão. Como a função aproximadora se mostra um método mais fácil e rápido de ser implementado, e por apresentar precisão igual ao processo numérico, optou-se por sua utilização.

Os parâmetros da distribuição gama têm valores mínimo, médio e máximo de 0,32449, 1,16340 e 11,33372, respectivamente, para alfa, e de 0,43699, 26,82905 e 345,71672, respectivamente, para beta. Trata-se de valores importantes, pois, quando utilizados em outras análises, no Estado de Minas Gerais, estes parâmetros terão uma faixa de abrangência próxima destes valores, norteando o pesquisador em seus resultados.

5.4.3. Distribuição log-normal

Embora estas distribuições sejam muito empregadas em outras áreas da análise climática e hidrológica, não apresentaram desempenho satisfatório na estimação das probabilidades, não sendo, em nenhum dos períodos considerados (conforme apresentam as Tabelas 9 e 10), uma boa distribuição para estimação dos dados diários e mensais observados.

5.4.4. Distribuição normal

Como citado anteriormente, há dois métodos para cálculo da função cumulativa de probabilidade da distribuição normal: a) pela resolução da equação de definição (Equação 37) utilizando método de integração de Gauss; ou b) usando uma função aproximadora (Equação 39).

Os dois métodos, em todos os cálculos apresentaram erro somente no 10^º dígito decimal, precisão esta que atende às exigências do presente trabalho.

Como a função aproximadora se mostra um método mais fácil e rápido de ser implementado, e por apresentar precisão igual ao processo de integração numérica, optou-se por sua utilização.

Ao observar os dados das Tabelas 11 a 18, pode-se notar que, para as estimativas diárias, decendiais e mensais, a aderência da normal é muito baixa; já para as estimativas totais, decendiais e mensais ela é alta. Ao analisar a distribuição de classes de freqüência de algumas estações isoladamente, nota-se que para os valores diários há maior freqüência nas classes iniciais, reduzindo bruscamente a partir da segunda ou terceira classe. Essa queda brusca não é acompanhada pelo modelo matemático da distribuição normal, como são os modelos da distribuição gama e Weibull, superestimando os valores das classes seguintes. Quando da aplicação dos testes de aderência, estas classes são reprovadas. Para os valores totais, decendiais e mensais, não ocorre variação brusca entre as classes iniciais, o que torna mais possível ao modelo sua estimação, aumentando sua aderência nesse período.

5.4.5. Distribuição Weibull

Os parâmetros da distribuição Weibull têm valores mínimo, médio e máximo de 0,10000, 2,57042 e 160,80000, respectivamente, para alfa, de 0,55191, 37,56826 e 482,87709, respectivamente, para beta, e de 0,44177, 1,04472 e 4,37132, respectivamente, para gama.

A distribuição Weibull é utilizada em análise hidrológica para eventos extremos; contudo, sua utilização em séries climáticas é pouco conhecida. No tópico seguinte será verificado que, para os valores diários de precipitação no período com chuva, esta distribuição se mostra um modelo dos mais precisos, superando a gama, que é uma distribuição muito empregada nestes casos.

5.5. Testes de aderência

Nas Tabelas 11 a 26 é apresentado o número de aderências aos testes de qui-quadrado e Kolmogorov-Smirnov, para os modelos estudados e períodos considerados.

O teste de qui-quadrado é específico para dados agrupados, em que as classes que possuem valores menores que três ou cinco devem ser agrupadas em outras classes, sendo um fator limitante para uso em dados com poucas classes. Este teste é baseado na soma dos erros absolutos das freqüências, que é comparada com um valor tabelado de acordo com o nível de significância desejado e os graus de liberdade da distribuição.

O teste de Kolmogorov-Smirnov pode ser usado tanto para dados agrupados quanto para dados individuais. Nos dados agrupados não há restrição quanto ao número nem ao valor das classes. É baseado no módulo da maior diferença entre a probabilidade observada e a estimada, que é comparada com um valor tabelado de acordo com o número de observações da série sob teste.

O teste de qui-quadrado é considerado mais rigoroso do que o de Kolmogorov-Smirnov, o que se deve, entre outros, aos seguintes aspectos:

- a) Considerando que uma distribuição sob teste tenha duas ou mais classes com probabilidades observadas diferentes das estimadas e, conseqüentemente, freqüências da mesma forma, quando se aplicam esses valores à equação de definição do teste qui-quadrado (Equação 52), tem-se uma somatória dos erros absolutos; aplicando a equação de definição do teste de Kolmogorov-Smirnov (Equação 55), tem-se um único valor, o módulo da diferença. Isto mostra que os erros, no teste de qui-quadrado, são considerados de forma cumulativa e em todas as classes e que, no teste de Kolmogorov-Smirnov, eles são considerados somente na classe em que foi maior.
- b) O quadro de freqüências de uma série pluviométrica mostra como estas são maiores nas classes iniciais e menores nas finais; a definição do teste de qui-quadrado determina que devem ser reunidas em uma única classe as classes com freqüências estimadas inferiores a três ou cinco. Os modelos testados geralmente superestimam as classes iniciais e subestimam as

finais, com algumas exceções. Para atender as definições do teste de qui-quadrado, as classes estimadas com frequência inferior a três ou cinco devem ser somadas a outra classe mais próxima. Decorre daí que a somatória ocorrerá também nas classes de frequência observadas, gerando um erro absoluto grande, que, somado aos anteriores, resulta em valores de qui-quadrado maiores que os tabelados, não aprovando a distribuição sob teste.

- c) Para o teste de Kolmogorov-Smirnov, com base na situação apresentada na letra anterior, não surgirá problema algum, pois, independentemente da distribuição de classes, o que interessa é o módulo da maior diferença, permitindo que o teste aprove a maioria das distribuições.
- d) Os valores de qui-quadrado calculados são comparados com valores críticos tabelados. Estes valores críticos são obtidos de tabelas referenciadas pelo nível de significância e pelo grau de liberdade, no caso do qui-quadrado, e pelo nível de significância e pelo número de observações, no teste de Kolmogorov-Smirnov. Observa-se que, neste teste, independentemente da capacidade da distribuição em estimar as frequências observadas e do número de classes, o valor crítico tabelado depende unicamente do número de observações, algo que não varia de distribuição para distribuição, dependendo apenas da série sob teste. Considerando agora o teste de qui-quadrado, vê-se que o grau de liberdade depende dos parâmetros da distribuição, em torno de dois ou três, e do número de classes. Analisando a observação feita na letra "a", este número reduz quando a distribuição subestima as classes finais, devido ao agrupamento de algumas classes em outras, e o grau de liberdade fica menor, reduzindo o valor crítico tabelado; isto mostra que o valor crítico para o qui-quadrado depende da capacidade da distribuição em estimar as frequências observadas, o que não ocorre no Kolmogorov-Smirnov.

O teste de Kolmogorov-Smirnov é um dos mais utilizados para análise de aderências de distribuições em estudos climáticos; contudo, o seu nível de aprovação de uma distribuição sob teste é muito elevado, como pode ser visto nas Tabelas 19 a 26, o que gera certa insegurança em relação aos critérios do teste, conforme exposto anteriormente. Tendo em vista as características

apresentadas pelo teste de qui-quadrado, optou-se pela sua escolha como referência em análises realizadas neste trabalho.

A Tabela 9 e 10 apresenta um resumo do desempenho das distribuições para os testes de qui-quadrado e Kolmogorov-Smirnov, considerando os meses do ano e períodos estudados. Observa-se que foi selecionada a distribuição que teve maior aderência entre as seis testadas, para cada período. Havendo empate, prevaleceu a que estava tendo melhor desempenho no contexto geral do período.

No caso das estimativas diárias de probabilidade, detectou-se a superioridade do desempenho da distribuição Weibull, com exceção dos decêndios do período, em que predomina a distribuição exponencial.

Nas estimativas totais de probabilidade para o período seco, é predominante a utilização da distribuição exponencial. Já no período com chuva há variação entre as distribuições Weibull, exponencial, gama e normal, nesta mesma ordem; a última aparece somente em dois meses. Em alguns casos, se fosse escolhida a segunda melhor distribuição para o período, sem distanciar muito da melhor, ter-se-ia uma menor variação no tipo de distribuição para o contexto dos valores mensais.

Tabela 9 - Resumo do ajustamento da funções de densidade de probabilidade estudadas, com base no teste de qui-quadrado para mês e período considerados (W= Weibull; E= exponencial; G= gama; N= normal)

Mês	Diário do Decêndio			Diário Mensal	Total do Decêndio			Total Mensal
	1º	2º	3º		1º	2º	3º	
janeiro	W	W	W	W	G	W	W	G
fevereiro	W	W	W	W	E	E	E	G
março	W	W	W	W	W	E	E	G
abril	W	W	W	W	E	E	E	G
maio	E	E	E	W	E	E	E	E
junho	E	E	E	W	E	E	E	E
julho	E	E	E	W	E	E	E	E
agosto	E	E	E	W	E	E	E	E
setembro	E	E	E	W	E	E	E	E
outubro	W	W	W	W	E	E	W	N
novembro	W	W	W	W	W	W	W	N
dezembro	W	W	W	W	G	G	W	G

Tabela 10 - Resumo do ajustamento das funções de densidade de probabilidade estudadas, com base no teste de Kolmogorov-Smirnov para mês e período considerados (W= Weibull; E= exponencial; G= gama; Ln=Lognormal; N= Normal)

Mês	Diário do Decêndio			Diário Mensal	Total do Decêndio			Total Mensal
	1º	2º	3º		1º	2º	3º	
janeiro	G	G	G	W	G	G	G	G
fevereiro	G	G	G	G	W	G	G	G
março	G	G	G	W	G	G	G	N
abril	G	G	W	W	G	G	G	G
maio	G	G	G	G	G	G	G	G
junho	G	Ln	G	G	G	G	G	G
julho	G	G	G	G	G	G	G	G
agosto	G	G	G	G	G	G	G	G
setembro	G	G	G	G	G	G	G	G
outubro	G	W	G	W	G	G	G	G
novembro	G	G	G	W	G	G	G	G
dezembro	G	G	G	G	G	G	G	G

Comparando as Tabelas 9 e 10, nota-se claramente que, para o teste de Kolmogorov-Smirnov, a distribuição gama tem grande aderência em todos os períodos considerados, o que não ocorre para o teste de qui-quadrado, sendo a distribuição Exponencial e Weibull as de melhor desempenho. Verificando as Tabelas 19 a 26 (teste de Kolmogorov-Smirnov), nota-se que ocorre o mesmo nível de aderência para mais de uma distribuição, o que não se observa nas Tabelas 11 a 18 (teste de qui-quadrado); isto mostra que a Tabela 10 poderia ter outra configuração de melhor distribuição, dependendo da afinidade do pesquisador por uma distribuição.

Tabela 11 - Número de aderências ao teste de qui-quadrado (Diário - 1º decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	76	72	64	56	61	55	48	31	71	71	69	75
Gama	84	80	79	74	55	41	29	17	60	69	71	76
LogNormal	29	40	35	52	30	22	11	8	27	43	41	25
LogNormal3	19	24	24	25	7	1	2	0	6	21	22	18
Normal	0	0	1	2	8	7	5	8	14	1	0	0
Weibull	86	82	87	80	53	40	27	19	60	72	77	78

Tabela 12 - Número de aderências ao teste de qui-quadrado (Diário - 2º decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	77	70	65	64	64	51	51	57	61	67	75	73
Gama	76	78	72	70	55	33	28	35	60	74	81	75
LogNormal	26	33	35	43	26	12	13	18	31	43	33	23
LogNormal3	17	19	19	14	6	2	1	1	9	25	25	16
Normal	0	0	0	3	9	10	7	12	9	1	0	0
Weibull	79	79	77	75	51	32	25	36	58	80	83	76

Tabela 13 - Número de aderências ao teste de qui-quadrado (Diário - 3º decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	74	73	56	62	54	53	47	55	75	64	69	69
Gama	78	80	72	60	45	29	32	43	75	70	78	74
LogNormal	26	44	40	44	37	13	18	23	33	39	29	24
LogNormal3	21	25	27	8	6	0	2	2	11	28	21	15
Normal	0	1	0	5	4	8	10	6	7	0	0	0
Weibull	85	82	77	64	50	28	28	44	75	76	81	77

Tabela 14 - Número de aderências ao teste de qui-quadrado (Diário mensal)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	50	47	35	28	49	47	54	52	57	38	46	54
Gama	67	67	60	53	52	56	58	52	69	57	60	67
LogNormal	4	15	17	39	43	37	41	37	31	19	10	5
LogNormal3	3	10	13	25	17	14	13	12	17	14	7	4
Normal	0	0	0	0	2	1	1	5	1	0	0	0
Weibull	75	74	73	70	65	61	59	56	73	70	67	71

Tabela 15 - Número de aderências ao teste de qui-quadrado (Total - 1º decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	58	82	72	83	55	32	22	5	39	75	61	60
Gama	67	63	76	61	39	11	9	2	26	61	75	83
LogNormal	33	27	31	26	12	4	3	0	5	21	33	45
LogNormal3	7	3	8	2	0	0	0	0	0	3	11	15
Normal	60	51	64	34	25	8	4	1	20	47	72	73
Weibull	65	67	76	64	39	10	8	1	26	63	79	82

Tabela 16 - Número de aderências ao teste de qui-quadrado (Total - 2º decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	63	78	75	73	53	11	11	11	41	78	63	40
Gama	76	71	69	58	26	3	5	1	24	74	72	76
LogNormal	31	27	24	23	9	1	1	0	6	27	28	41
LogNormal3	10	8	5	1	0	0	0	0	0	5	8	16
Normal	74	66	61	38	18	1	3	2	12	55	75	75
Weibull	80	77	74	58	27	3	5	1	23	74	79	74

Tabela 17 - Número de aderências ao teste de qui-quadrado (Total - 3º decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	71	71	79	66	49	13	18	17	63	67	54	58
Gama	73	62	69	47	24	5	9	6	56	71	76	77
LogNormal	25	20	25	19	6	0	3	1	19	25	45	39
LogNormal3	7	4	6	1	0	0	0	0	2	5	12	14
Normal	70	47	48	28	12	3	7	5	46	70	71	75
Weibull	75	63	67	45	24	5	9	4	56	75	78	78

Tabela 18 - Número de aderências ao teste de qui-quadrado (Total mensal)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	31	33	28	61	80	68	58	56	79	29	4	4
Gama	82	84	82	83	68	48	40	33	72	77	80	86
LogNormal	54	42	46	57	36	18	13	11	30	47	58	71
LogNormal3	12	6	9	12	7	1	1	0	2	7	8	7
Normal	82	82	82	71	51	25	28	16	56	82	83	83
Weibull	59	74	62	78	71	47	42	33	75	66	45	31

Tabela 19 - Número de aderências ao teste de Kolmogorov-Smirnov (Diário - 1º decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	97	93	96	97	99	98	93	92	99	98	95	94
Gama	100	100	100	100	100	99	95	92	99	100	99	99
LogNormal	75	84	85	94	98	97	94	92	96	98	83	74
LogNormal3	35	42	44	62	68	71	69	71	75	58	45	31
Normal	1	2	3	8	33	49	53	65	49	11	2	1
Weibull	99	99	99	100	100	98	95	89	99	99	99	99

Tabela 20 - Número de aderências ao teste de Kolmogorov-Smirnov (Diário - 2º decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	97	96	97	95	98	96	97	96	98	92	96	97
Gama	99	99	99	99	100	96	98	97	100	98	99	99
LogNormal	73	83	88	95	98	97	97	97	97	87	81	68
LogNormal3	32	39	50	53	59	67	71	76	70	48	36	25
Normal	2	1	1	11	41	57	61	68	40	3	2	0
Weibull	97	99	99	99	99	96	97	97	99	99	99	99

Tabela 21 - Número de aderências ao teste de Kolmogorov-Smirnov (Diário - 3^o decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	94	96	96	96	99	96	100	98	98	96	94	96
Gama	99	100	100	99	100	96	100	99	100	98	100	100
LogNormal	73	91	91	99	98	96	99	99	92	85	83	65
LogNormal3	37	51	46	67	70	68	75	81	56	43	39	28
Normal	0	7	2	19	22	50	58	53	23	3	0	1
Weibull	98	99	99	100	100	96	99	98	100	98	99	99

Tabela 22 - Número de aderências ao teste de Kolmogorov-Smirnov (Diário mensal)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	83	84	73	78	93	96	94	97	95	79	80	81
Gama	93	96	94	93	99	99	98	98	99	92	92	94
LogNormal	27	46	50	86	93	93	95	97	76	62	40	26
LogNormal3	6	5	8	17	30	38	43	50	24	10	3	2
Normal	0	0	0	1	3	14	14	18	3	0	0	0
Weibull	95	95	97	96	99	99	98	98	98	94	94	92

Tabela 23 - Número de aderências ao teste de Kolmogorov-Smirnov (Total - 1^o decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	84	97	92	100	100	99	94	91	100	99	86	81
Gama	99	99	100	100	100	99	94	91	100	100	100	100
LogNormal	91	92	93	99	100	99	93	91	99	96	93	95
LogNormal3	60	78	75	89	92	91	87	83	92	84	64	67
Normal	93	90	95	81	89	85	80	83	89	89	96	98
Weibull	91	100	98	100	99	99	94	90	100	98	97	98

Tabela 24 - Número de aderências ao teste de Kolmogorov-Smirnov (Total - 2º decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	84	95	96	100	100	96	98	95	100	98	90	66
Gama	100	100	100	100	100	96	98	95	100	100	100	100
LogNormal	90	93	93	99	100	96	98	95	100	97	90	97
LogNormal3	72	76	78	88	88	92	92	92	95	80	67	53
Normal	96	96	93	86	80	79	86	86	89	89	96	96
Weibull	99	100	100	99	100	96	97	95	99	99	99	94

Tabela 25 - Número de aderências ao teste de Kolmogorov-Smirnov (Total - 3º decêndio)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	93	97	99	100	100	95	98	99	99	91	77	80
Gama	100	100	100	100	100	95	99	99	99	100	100	99
LogNormal	93	93	97	99	100	95	99	99	97	96	94	94
LogNormal3	72	82	82	87	94	88	91	92	75	75	46	60
Normal	96	88	95	83	84	81	92	80	91	96	97	96
Weibull	100	99	99	100	100	95	98	98	99	99	95	98

Tabela 26 - Número de aderências ao teste de Kolmogorov-Smirnov (Total mensal)

Distribuição/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Exponencial	42	49	44	75	99	100	99	99	98	50	14	10
Gama	100	100	98	99	100	100	100	99	99	100	100	99
LogNormal	98	92	92	96	98	99	100	99	97	96	92	95
LogNormal3	28	43	37	55	77	92	89	89	85	37	15	8
Normal	98	97	100	96	85	79	82	77	90	97	98	99
Weibull	74	88	80	89	99	99	100	99	99	83	56	38

5.6. Análise de agrupamento

Com as variáveis para análise de agrupamento (Tabela 7) obtidas das 624 estações, foram feitas combinações entre os métodos de agrupamento (encadeamento simples e encadeamento completo) e as medidas de

dissimilaridade (distância euclidiana, distância euclidiana média e distância de Mahalanobis), aplicados aos dados de precipitação, considerando também a otimização pelo método de Toucher.

Em relação às combinações anteriores, a que apresentou melhor distribuição de grupos foi o processo de encadeamento completo com distância euclidiana, cujo dendograma pode ser visto na Figura 8. Este dendograma não apresenta grandes saltos entre os maiores e os menores grupos, característica da distância euclidiana média. Não se observa também grande número de grupos, característica do encadeamento simples.

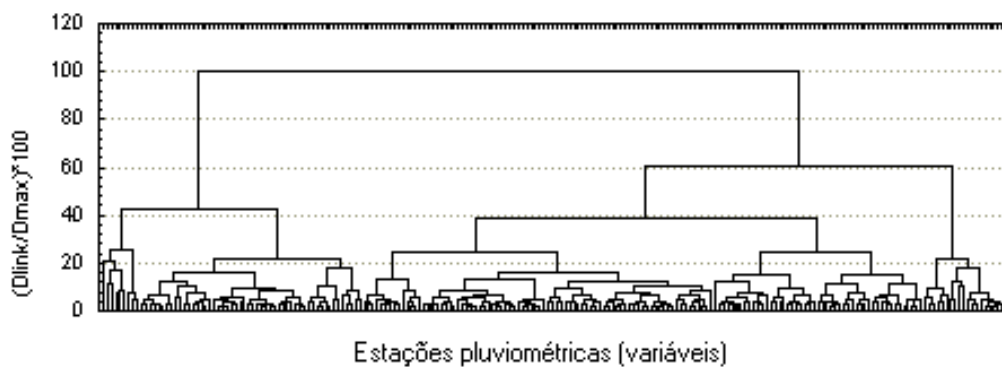


Figura 8 Dendograma para 212 estações, agrupamento por encadeamento completo e distância euclidiana.

A determinação do número de grupos para o dendograma é feita de forma visual, e o índice de salto escolhido foi de aproximadamente 25%, perfazendo um total de sete grupos. Índices de salto maiores iriam reduzir muito o número de grupos, podendo omitir características importantes de algum outro subgrupo. Para índices de salto menores, o número de agrupamentos é grande dificultando a análise, pois haveria a sobreposição geográfica de grupos e muitos grupos com poucas estações.

A Tabela 2A apresenta o código das estações de cada grupo. A Figura 9 mostra a disposição geográfica das estações por grupo. Foi omitido o grupo 7, pois este apresentava somente duas estações no sul da região sudoeste do

Estado. As Tabelas 27, 28 e 29 apresentam os valores mínimo, médio e máximo de precipitação para cada grupo, respectivamente.

Observa-se, na Figura 9, que existem grupos de estações que se sobrepõem geograficamente; reportando ao dendograma (Figura 8), pode-se verificar que os desníveis de salto são quase imperceptíveis dentro de um subgrupo, o que se deve à capacidade do método em separar estações com pequenas diferenças de características; consideram-se também o número e a continuidade de dados de cada estação, uma vez que a mesma região possui estações com grande diversidade de dados, influenciando a variância e a média.

É de se esperar que as estações de uma mesma região pertençam ao mesmo grupo, mas pode-se observar que isto não ocorre, pois há dois grupos em uma mesma região, como os grupos 1 e 3, e da mesma forma o grupo 4 e 5, devendo-se isto aos microclimas locais, que proporcionam valores diferentes para estações que pertencem a uma mesma região climática.

Outro fato importante é a existência de estações em regiões diferentes que pertencem ao mesmo grupo, como os grupos 1, 4, 5 e 6. Isto se deve também aos microclimas locais, que, mesmo em regiões de climas diferentes, podem apresentar a mesma característica.

A análise de agrupamento é uma ferramenta importante para se concluir que em uma mesma região podem existir estações pluviométricas com totais médios diferentes, formando vários subgrupos. Ela fornece também informações sobre a influência de um microclima na análise meteorológica, evitando o conceito da generalização de informações por região.

Com os dados das Tabelas 27, 28 e 29 tem-se uma noção dos totais médios mensais de precipitação para cada grupo. Se comparados com os gráficos das Figuras 3 e 4, vê-se uma correspondência de valores para todos os meses e que, para todos os grupos, os meses de janeiro, fevereiro, março, outubro, novembro e dezembro apresentam precipitação média acima de 100 mm; e para os meses de abril, maio, junho, julho, agosto e setembro, as médias ficam abaixo de 50 mm, com poucas exceções.

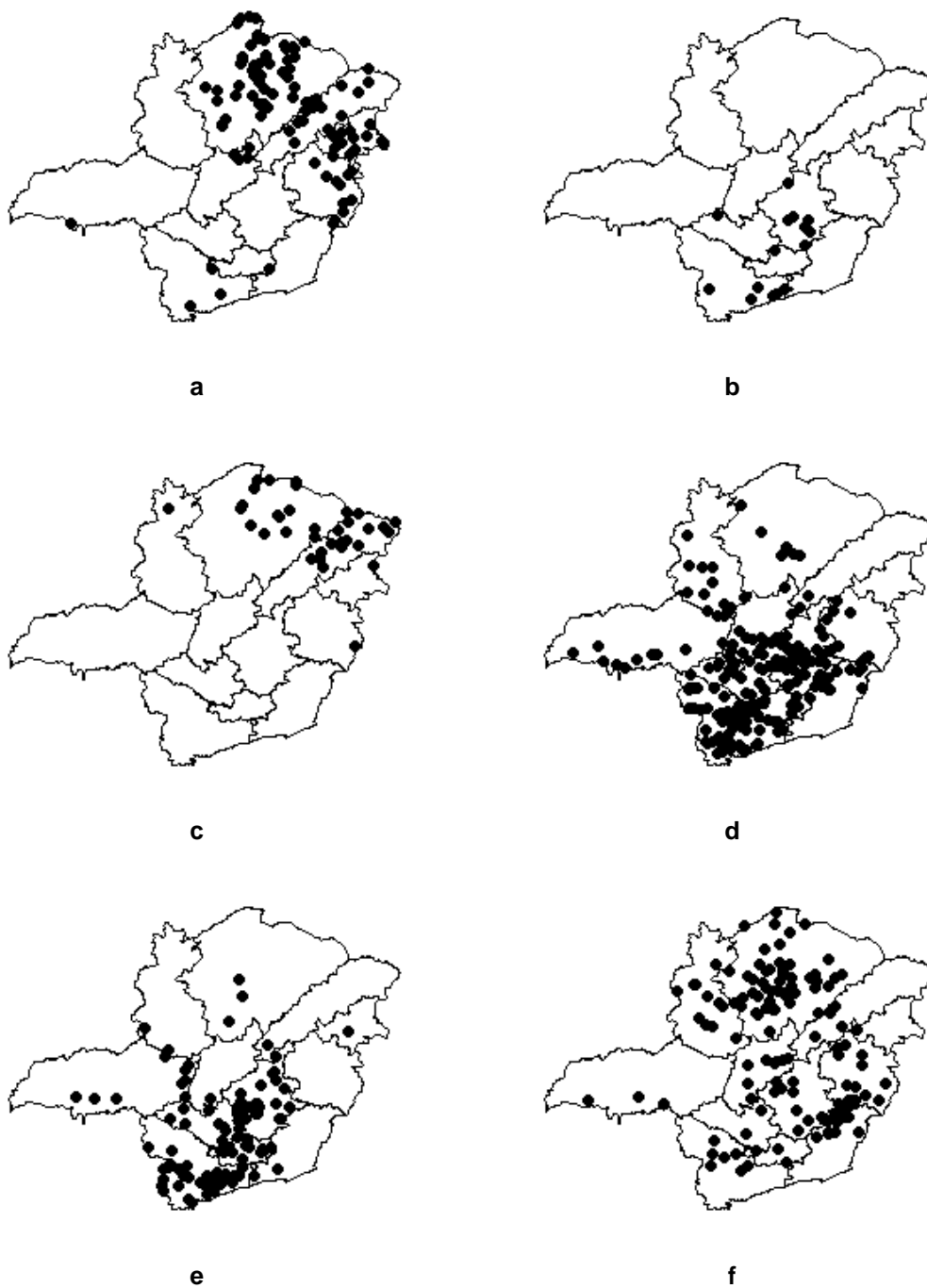


Figura 9 - Disposição geográfica das estações por grupo no Estado de Minas Gerais: a) Grupo 1; b) Grupo 2; c) Grupo 3; d) Grupo 4; e) Grupo 5; f) Grupo 6.

Tabela 27 - Precipitação mínima mensal de janeiro a dezembro, por grupo

Grupo/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Grupo 1	82	68	55	23	0	0	0	0	6	48	96	126
Grupo 2	270	200	164	61	27	7	11	7	25	110	189	286
Grupo 3	68	26	49	18	3	0	0	0	7	34	95	99
Grupo 4	154	116	102	40	7	0	2	1	13	72	126	183
Grupo 5	226	144	110	45	16	5	1	3	15	86	145	236
Grupo 6	121	91	50	31	7	0	1	0	8	60	126	176
Grupo 7	395	345	285	79	66	39	23	27	64	150	226	330

Tabela 28 - Precipitação média mensal de janeiro a dezembro, por grupo

Grupo/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Grupo 1	168	108	111	54	18	10	11	8	23	88	172	199
Grupo 2	344	256	221	88	47	29	23	25	68	160	245	377
Grupo 3	136	79	87	43	17	8	10	7	19	66	145	163
Grupo 4	257	181	159	66	34	17	14	14	48	116	189	270
Grupo 5	293	204	178	76	42	22	18	19	58	130	207	293
Grupo 6	217	142	136	61	22	9	9	8	30	102	189	236
Grupo 7	417	352	297	93	69	40	27	29	65	159	243	338

Tabela 29 - Precipitação máxima mensal de janeiro a dezembro, por grupo

Grupo/Mês	Jan.	Fev.	Mar.	Abr.	Mai	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Grupo 1	237	165	187	102	81	102	113	92	65	151	318	329
Grupo 2	412	338	276	128	64	75	67	72	122	231	300	565
Grupo 3	194	125	121	90	57	54	57	34	39	109	211	242
Grupo 4	335	303	248	108	97	44	39	41	92	187	333	398
Grupo 5	389	298	237	120	85	56	47	37	104	180	289	385
Grupo 6	303	227	222	110	50	29	31	25	73	153	329	330
Grupo 7	438	359	309	107	72	40	30	31	67	169	259	346

Necessita-se de mais estudos sobre a aplicação de técnicas de análise de agrupamento a uma grande quantidade de estações climatológicas, uma vez que, como pode ser visto no trabalho de RIBEIRO et al. (1996), para pequeno número de estações, a formação de grupos foi bem coerente, não

havendo sobreposição de grupos. Isto não aconteceu no presente trabalho, não permitindo uma distribuição homogênea para uma região, pois havia a sobreposição de grupos, devido à grande quantidade de estações consideradas e ao fato de os valores médios serem próximos.

6. CONCLUSÕES

Considerando os resultados apresentados, conclui-se que:

- Para as estimativas diárias de probabilidade, destaca-se o desempenho da distribuição Weibull, com exceção dos decêndios do período seco, em que predomina a distribuição exponencial.
- Nas estimativas mensais de probabilidade para o período seco, é predominante a utilização da distribuição exponencial, e, para o período com chuva, há variação entre as distribuições Weibull, exponencial, gama e normal, nesta ordem, com esta última aparecendo somente em dois meses.
- O teste de qui-quadrado apresentou melhores características para verificar o ajustamento de uma distribuição de probabilidade estimada a dados observados.
- Tanto para o cálculo da função cumulativa de probabilidade da distribuição gama quanto para o da normal, o emprego de uma função aproximadora se mostrou com precisão aceitável e de fácil implementação, em vez da utilização da equação de definição com o cálculo numérico da integral.
- O cálculo da função gama pela equação polinomial (Equação 21) se mostrou mais preciso do que os outros métodos citados.
- A análise de agrupamento é uma ferramenta importante para se concluir que, em uma mesma região, podem existir estações

pluviométricas com totais médios diferentes, evitando a generalização de informações por região.

- A medida de dissimilaridade distância euclidiana e o método de agrupamento encadeamento completo, apresentaram uma formação de grupos mais coerente que as outras combinações testadas para a os dados em questão.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAMOWITZ, M., STEGUN, I. A. **Handbook of mathematical functions**, New York, Dover, 1972, 1046p.
- ALMEIDA, R. M. B. **Características climatológicas do regime de chuva em Minas Gerais**, Tese de Mestrado, Viçosa, UFV, 1995, 64p.
- ASSAD, E. D. **Chuvvas nos cerrados**, Embrapa-DF, 1994, 423p.
- ASSIS, F. N. DE., ARRUDA, H.V., PEREIRA, A. R. **Aplicações estatísticas a climatologia**, Pelotas, Ed. Universitária/UFPel, 1996, 161p.
- ASSIS, S. V. de **Modelo para estimar o número de dias trabalháveis com tratores agrícolas em função de parâmetros do solo e do clima**, Viçosa, UFV, 1988, 92p.
- BELTRAME, L. F. S. **Probabilidade de ocorrência de deficit e excesso hídrico em solos do Rio Grande do Sul**, Tese de Mestrado, Porto Alegre, UFRGS, 1979, 79p, il.
- BURK, THOMAS E.; NEWBERRY, JAMES D.A **Simple algorithm for moment-based recovery weibull distribution**, Forest Science, v30, no 2, 1984, p329-332.
- CAMPOS, H. **Estatística experimental não paramétrica**, Piracicaba, São Paulo, ESALQ/USP, 1979, 3 ed, 200p.
- CARMARGO, MARCELO B P et. al. **Spatial and temporal variability of daily weather variables in sub-humid and semi-arid area of the united states high plains**, Agricultural and Forest meteorology, v93, 1999, p141-148

- CHATFIELD, C. **Proximate composition of american food materials**, Washington, US, 1940, 90p, il.
- COUTO, H. T. Z. **Distribuições de diâmetro em plantações de *pinus caribaea morelet***, Tese de Mestrado, ESALQ, 1980, 79p.
- CRUZ, C. D., REGAZZI, A. J. **Modelos biométricos aplicados ao melhoramento genético**, 2 edição, Viçosa/MG, UFV, 1997, 390p.
- CUPOLILLO, F. **Períodos de estiagem durante a estação chuvosa no estado de Minas Gerais: espacialização e aspectos dinâmicos relacionados**, Viçosa, UFV, 1997, 148p.
- DUAN, JINFAN et. al. **Evaluation of probability density function in precipitation models for the pacific northwest**, Journal of the American Water Resources Association, v34, n^o3, 1998, p617627
- EVERITTI, B. S. **Cluster analysis**, New York, 1980, 136p.
- FARIA, R. A. **Demanda de irrigação suplementar no estado de Minas Gerais**, Viçosa, UFV, 1998, 75p.
- FRANCO, J. C. G. **Erosividade da chuva - distribuição e relações com a precipitação em Porto Presidente Stroessner e Villa Rica Paraguai**, ESALQ, SP, 1986, 110p.
- GOVE, JEFFREY H. et. al. **Maximum-Likelihood estimation of Weibull function parameters using a general interactive optimizer and grouped data**, Forest Ecology and Management, Elsevier Science Publishers, v28, 1989, p61-69
- GREENWOOD, J. A.; DURAND, D. **Aids for fitting the gamma distribution by maximum likelihood**, Technometrics, v2, n1, 1960, p55-65.
- HAAG, H. P. **Chuvas ácidas**, Campinas, São Paulo, Fundação Cargil, 1985, 77p, il.
- HAAN, C. T. **Statistical methods in hidrology**, Ames, Iowa State University Press, 1977, 378p, il.
- HASTINGS, N. A. J., PEACOCK, J. B. **Statistical distributions: A handbook for students and practitioners**, Longon Butterworths, England, 1975, 129p.

- HUF, F. A. e NEIL, J.C. **Comparison of several methods for rainfall frequency analysis**. Illinois State Water Survey Urbana, Illinois. Journal of Geophysical Research, 64(5):541-547, 1959.
- HUSCHKE, R. E. **Glossary of meteorology**, Boston, American Meteorology Society, 1970, 638p.
- JOHNSON, N. L e KOTZ, S. **Distribution in statistics**, continuo univariate distribution, New York, Houghton Mifflin, 1970, 2v.
- KITE, G. W. **Frequency and risk analisys in hidrology**. Fort Colins, Water Resourses Publications, 1978, v3, 395p.
- KRUG, ALBERT G. et al. **Determining initial values for parameters of a Weibull model: a case study**, Forest Science, v30, nº 3, 1984, p573-581.
- MIRSHAWKA, V. **Estatística**, Vol II, São Paulo, Liv. Novel, 1971, 367p.
- PACITTI, T. **Fortran**. Rio de Janeiro, Livro Técnico e Científico Editora, 1974, 377p.
- PAIS DE MELO, J. S. **Modelo versátil para estimar as probabilidades de dias chuvosos em intervalos mensais e de alturas diárias de chuvas**, Viçosa, UFV, 1989, 93p.
- RAO, C. R. **Advanced statistical methods in biometric research**, New York, John Wiley, 1952, 390p.
- RIBEIRO, C. R et. al. . **Análise do regime pluviométrico da região amazônica a partir de dados de onze localidades**, Revista Brasileira de meteorologia, 11(1/2), 1996, p25-35.
- SEDIYAMA, G. C., Nota de aula, 1999.
- THOM, H.C.S. **A note on the gamma distribution**. Monthly Weather Review, Washington, v86, 1958, p.117-122.
- TOMMASELLI, J. T. G. **Deficiência hídrica no solo e épocas de plantio de milho em Cambara e Londrina PR**, Piracicaba, ESALQ, 1992, 78p, il.
- WMO - World Meteorological Organization - nota técnica nº 81.
- WOLF, FRANK L. **Elements of probability and statistical**, New York, mcGraw-Hill, 1962, 332p, il.

APÊNDICE

APÊNDICE

Tabela 1A - Identificação das 814 estações utilizadas na análise

Estações do INMET											
00083437; 00083442; 00083481; 00083488; 00083492; 00083531; 00083543; 00083577; 00083579; 00083582; 00083586; 00083587; 00083589; 00083592; 00083632; 00083642; 00083683; 00083687; 00083689; 00083692; 00083736											
Estações da Aneel											
01442000;	01442019;	01442020;	01442022;	01443000;	01443001;	01443003;	01443004;	01443007;	01443008;	01443009;	01443017;
01443020;	01444000;	01444001;	01444002;	01444003;	01444004;	01444005;	01445000;	01539035;	01540013;	01541002;	01541003;
01542000;	01542001;	01542002;	01542006;	01542016;	01543000;	01543001;	01543002;	01543003;	01543004;	01543005;	01543007;
01543008;	01543009;	01543010;	01543011;	01543013;	01544000;	01544001;	01544005;	01544006;	01544007;	01544009;	01544010;
01544011;	01544012;	01544013;	01544014;	01544015;	01544016;	01544017;	01544018;	01544019;	01544020;	01544021;	01544030;
01544031;	01545002;	01546000;	01546001;	01546005;	01547021;	01547022;	01639000;	01639004;	01639006;	01639010;	01639013;
01639015;	01639018;	01640000;	01640001;	01640002;	01640007;	01640008;	01641000;	01641001;	01641002;	01641003;	01641004;
01641007;	01641008;	01641010;	01641011;	01641012;	01642000;	01642001;	01642002;	01642003;	01642004;	01642005;	01642006;
01642007;	01642008;	01642009;	01642010;	01642011;	01642012;	01642013;	01642014;	01642025;	01643001;	01643002;	01643003;
01643004;	01643005;	01643006;	01643007;	01643008;	01643010;	01643012;	01643013;	01643014;	01643015;	01643017;	01643018;
01643019;	01643020;	01644000;	01644001;	01644002;	01644004;	01644005;	01644006;	01644007;	01644009;	01644010;	01644012;
01644013;	01644014;	01644015;	01644016;	01644017;	01644018;	01644019;	01644020;	01644021;	01644022;	01644023;	01644024;
01644025;	01644027;	01644028;	01645000;	01645001;	01645002;	01645003;	01645005;	01645007;	01645009;	01645013;	01646000;
01646001;	01646003;	01646004;	01647008;	01740000;	01740001;	01740002;	01740003;	01740005;	01740006;	01740007;	01740008;
01740019;	01740020;	01740021;	01740022;	01740026;	01741001;	01741003;	01741006;	01741007;	01741008;	01741009;	01741013;
01742000;	01742001;	01742002;	01742003;	01742006;	01742008;	01742012;	01742013;	01742014;	01742015;	01742017;	01742020;
01743000;	01743002;	01744000;	01744001;	01744002;	01744004;	01744005;	01744006;	01744008;	01744009;	01744010;	01744011;
01744012;	01744014;	01744025;	01744030;	01745000;	01745001;	01745007;	01746001;	01746002;	01746004;	01746006;	01746007;
01746008;	01746012;	01747005;	01840000;	01840004;	01840006;	01840007;	01840009;	01840010;	01840011;	01840012;	01840013;
01840014;	01840015;	01840017;	01840021;	01841001;	01841003;	01841005;	01841006;	01841007;	01841008;	01841009;	01841010;
01841011;	01841014;	01841015;	01841018;	01842000;	01842002;	01842003;	01842004;	01842005;	01842006;	01842007;	01842008;
01843000;	01843001;	01843002;	01843003;	01843005;	01843009;	01844000;	01844001;	01844002;	01844003;	01844004;	01844009;
01844010;	01844015;	01844018;	01845000;	01845002;	01845004;	01845008;	01845009;	01845010;	01845013;	01845014;	01845021;
01846003;	01846005;	01846015;	01846016;	01846017;	01846018;	01940006;	01940008;	01940009;	01940011;	01940012;	01940013;
01940014;	01940016;	01940019;	01941000;	01941001;	01941002;	01941004;	01941005;	01941006;	01941007;	01941008;	01941009;
01941010;	01941011;	01941012;	01941019;	01942001;	01942002;	01942003;	01942005;	01942006;	01942007;	01942008;	01942009;
01942021;	01943000;	01943001;	01943002;	01943003;	01943004;	01943005;	01943006;	01943007;	01943008;	01943009;	01943010;
01943011;	01943012;	01943013;	01943017;	01943021;	01943022;	01943023;	01943024;	01943025;	01943026;	01943027;	01943028;
01943029;	01943030;	01943031;	01943032;	01943033;	01943034;	01943035;	01943039;	01943042;	01943049;	01943053;	01943055;
01943061;	01944000;	01944003;	01944004;	01944005;	01944007;	01944009;	01944010;	01944011;	01944014;	01944015;	01944016;
01944018;	01944019;	01944020;	01944021;	01944022;	01944023;	01944024;	01944026;	01944027;	01944029;	01944031;	01944032;
01944033;	01944040;	01944048;	01944049;	01944051;	01944052;	01944055;	01944059;	01945000;	01945002;	01945004;	01945008;
01945012;	01945013;	01945014;	01945015;	01945016;	01945019;	01945020;	01945035;	01945037;	01945038;	01945039;	01946000;
01946009;	01947000;	01947002;	01947016;	01947020;	01948000;	01948001;	01948003;	01948007;	01949000;	01949001;	01949004;
01949005;	01949007;	01950000;	01950001;	01950002;	01950003;	01950005;	02041005;	02041007;	02041008;	02041009;	02041012;
02041013;	02041016;	02041017;	02041018;	02041019;	02041020;	02041022;	02041023;	02041024;	02041030;	02042000;	02042001;
02042002;	02042004;	02042007;	02042008;	02042009;	02042010;	02042011;	02042012;	02042015;	02042016;	02042017;	02042018;
02042022;	02042024;	02042031;	02043002;	02043003;	02043004;	02043005;	02043007;	02043008;	02043009;	02043010;	02043011;
02043013;	02043014;	02043016;	02043017;	02043018;	02043019;	02043022;	02043023;	02043024;	02043025;	02043026;	02043027;
02043028;	02043031;	02043042;	02043043;	02043044;	02043057;	02043058;	02043059;	02044000;	02044001;	02044002;	02044003;
02044005;	02044006;	02044007;	02044008;	02044009;	02044012;	02044014;	02044016;	02044017;	02044018;	02044019;	02044020;
02044021;	02044022;	02044023;	02044026;	02044027;	02044036;	02044037;	02044038;	02044039;	02044040;	02044041;	02044042;
02044043;	02044046;	02044047;	02044049;	02044050;	02044052;	02044053;	02045000;	02045001;	02045002;	02045004;	02045005;
02045009;	02045010;	02045011;	02045012;	02045013;	02045015;	02045020;	02045021;	02045026;	02046000;	02046001;	02046002;
02046004;	02046005;	02046007;	02046009;	02046011;	02046013;	02046015;	02046025;	02047001;	02047002;	02047004;	02047005;
02047006;	02047007;	02047008;	02047009;	02047010;	02047011;	02047012;	02047013;	02047016;	02047017;	02047020;	02047026;
02047028;	02047031;	02047034;	02047057;	02047062;	02047063;	02047065;	02047067;	02047078;	02047080;	02047081;	02047082;
02047083;	02047091;	02047101;	02047102;	02048001;	02048002;	02048003;	02048004;	02048005;	02048006;	02048007;	02048013;
02048014;	02048041;	02048049;	02048072;	02048075;	02048076;	02048077;	02048078;	02048080;	02048091;	02049001;	02049002;
02049003;	02049012;	02049013;	02049014;	02049031;	02049033;	02049039;	02049042;	02049043;	02049047;	02049054;	02049057;
02050001;	02050002;	02050003;	02050004;	02050005;	02050006;	02050007;	02050025;	02050039;	02141016;	02143003;	02143004;
02143005;	02143006;	02143007;	02143008;	02143009;	02143010;	02143055;	02143062;	02144000;	02144001;	02144002;	02144003;
02144004;	02144005;	02144006;	02144007;	02144008;	02144009;	02144010;	02144011;	02144013;	02144014;	02144015;	02144016;
02144018;	02144019;	02144020;	02144021;	02144022;	02144023;	02144024;	02144025;	02144026;	02144037;	02144038;	02145000;
02145001;	02145002;	02145003;	02145004;	02145005;	02145006;	02145007;	02145008;	02145009;	02145010;	02145012;	02145013;
02145016;	02145017;	02145018;	02145020;	02145021;	02145022;	02145023;	02145024;	02145032;	02145033;	02145034;	02145036;
02145039;	02145040;	02145041;	02145042;	02145043;	02145044;	02146000;	02146001;	02146002;	02146003;	02146004;	02146005;
02146007;	02146009;	02146010;	02146012;	02146013;	02146014;	02146015;	02146016;	02146017;	02146019;	02146023;	02146024;
02146026;	02146027;	02146028;	02146029;	02146030;	02146036;	02146042;	02146043;	02146045;	02146048;	02146052;	02146053;
02146054;	02146055;	02146057;	02146080;	02146081;	02147001;	02147003;	02147023;	02147048;	02147054;	02147057;	02147069;
02147071;	02147078;	02147108;	02147165;	02147166;	02244054;	02244056;	02244057;	02244058;	02244062;	02244063;	02244064;
02244065;	02244068;	02244071;	02244075;	02245000;	02245005;	02245012;	02245019;	02245064;	02245065;	02245066;	02245067;
02245068;	02245069;	02245070;	02245071;	02245073;	02245074;	02245075;	02245077;	02245080;	02245081;	02245082;	02245083;
02245084;	02245085;	02245086;	02245087;	02245088;	02245089;	02245090;	02245094;	02245103;	02245104;	02245107;	02245112;
02245147;	02245167;	02245173;	02246001;	02246004;	02246006;	02246008;	02246011;	02246013;	02246014;	02246016;	02246017;
02246018;	02246047;	02246048;	02246049;	02246050;	02246051;	02246052;	02246055;	02246056;	02246065;	02246098;	02246127 e 02246131.

Nota: As demais informações, tais como latitude, longitude e outras, podem ser encontradas nos sites dos respectivos órgãos gerenciadores

Tabela 2A - Código das estações de cada grupo resultante da análise de agrupamento

Grupo	Código das estações
1	00083488; 00083492; 00083543; 01442020; 014430001; 01443003; 01443008; 01443009; 01444000; 01444001; 01444004; 01444005; 01540013; 01542000; 01542001; 01542002; 01543001; 01543002; 01543003; 01543004; 01543005; 01543007; 01543008; 01543009; 01543010; 01543011; 01544000; 01544001; 01544005; 01544007; 01544015; 01544017; 01544019; 01544030; 01544031; 01640001; 01641002; 01641011; 01642001; 01642002; 01642005; 01642006; 01642007; 01642011; 01642012; 01642013; 01643005; 01643007; 01643010; 01643014; 01643019; 01643020; 01644002; 01644006; 01644022; 01644027; 01645001; 01645002; 01645009; 01740000; 01740001; 01740002; 01740026; 01741001; 01741003; 01741006; 01741008; 01741009; 01741013; 01742000; 01742003; 01742008; 01742015; 01743000; 01744000; 01744005; 01744011; 01744025; 01744030; 01841003; 01841006; 01841007; 01841008; 01841011; 01841014; 01841018; 01842003; 01844001; 01844002; 01844018; 01941005; 01941006; 01941010; 01941019; 02041007; 02041008; 02049031; 02143062; 02145006; 02244063; 02245094
2	01943022; 01944020; 01946000; 02043003; 02043017; 02043022; 02043059; 02044000; 02044012; 02144003; 02244057; 02244058; 02244064; 02245080; 02246047
3	00083442; 01442019; 01442022; 01443000; 01443007; 01541002; 01541003; 01542016; 01543000; 01543013; 01544006; 01544010; 01544018; 01546001; 01639015; 01640000; 01640002; 01640007; 01641000; 01641001; 01641003; 01641004; 01641007; 01641008; 01641010; 01641012; 01642000; 01642008; 01642009; 01642025; 01643002; 01643012; 01644017; 01740003; 01742002; 01742012; 01742013; 01742020; 01941002
4	00083577; 00083579; 00083582; 00083586; 00083632; 01545002; 01643006; 01643013; 01644021; 01644024; 01644025; 01644028; 01646001; 01744012; 01746002; 01746004; 01746008; 01746012; 01842000; 01842005; 01842006; 01842007; 01843000; 01843002; 01843003; 01845002; 01845008; 01845009; 01845010; 01845013; 01846003; 01846015; 01846016; 01846017; 01941000; 01941007; 01942003; 01942005; 01942009; 01943001; 01943003; 01943004; 01943005; 01943007; 01943009; 01943010; 01943012; 01943013; 01943017; 01943021; 01943023; 01943024; 01943026; 01943028; 01943029; 01943030; 01943031; 01943032; 01943033; 01943034; 01943053; 01943061; 01944000; 01944004; 01944005; 01944007; 01944009; 01944010; 01944011; 01944014; 01944015; 01944016; 01944019; 01944021; 01944023; 01944024; 01944027; 01944031; 01944032; 01944033; 01944040; 01944048; 01944049; 01944052; 01945000; 01945002; 01945012; 01945014; 01945016; 01945019; 01945020; 01945035; 01945037; 01945038; 01947000; 01948001; 01949004; 01949007; 01950000; 02041005; 02041009; 02042002; 02042004; 02042007; 02042011; 02042018; 02042022; 02042024; 02043007; 02043009; 02043010; 02043011; 02043013; 02043014; 02043018; 02043019; 02043023; 02043025; 02043044; 02043057; 02043058; 02044002; 02044003; 02044005; 02044006; 02044007; 02044008; 02044018; 02044020; 02044027; 02044040; 02044047; 02044049; 02044053; 02045000; 02045002; 02045004; 02045009; 02045010; 02045011; 02045015; 02045021; 02045026; 02046000; 02046001; 02046002; 02046004; 02046007; 02046009; 02046011; 02046025; 02048041; 02048049; 02143004; 02143006; 02143008; 02143055; 02144000; 02144001; 02144004; 02144006; 02144008; 02144010; 02144019; 02144021; 02144026; 02144038; 02145000; 02145002; 02145003; 02145005; 02145007; 02145008; 02145009; 02145010; 02145017; 02145018; 02145021; 02145022; 02145023; 02145024; 02145032; 02145034; 02145036; 02145041; 02145042; 02145043; 02145044; 02146000; 02146024; 02146026; 02146030; 02146036; 02146080; 02146081; 02244056; 02244062; 02244068; 02245000; 02245064; 02245066; 02245070; 02245071; 02245073; 02245074; 02245077; 02245083; 02245084; 02245085; 02245086; 02246048; 02246049; 02246050; 02246056
5	00083531; 00083587; 00083589; 00083683; 00083687; 00083689; 00083692; 00083736; 01644009; 01644020; 01741007; 01744009; 01747005; 01843001; 01843009; 01845004; 01845014; 01846005; 01943000; 01943002; 01943008; 01943011; 01943025; 01943027; 01943035; 01943039; 01943055; 01944003; 01944022; 01944026; 01944051; 01944055; 01945008; 01945013; 01945015; 01946009; 01947016; 01948000; 01948007; 01949005; 02043002; 02043004; 02043008; 02043028; 02043031; 02043042; 02043043; 02044001; 02044009; 02044016; 02044017; 02044019; 02044021; 02044022; 02044023; 02044026; 02044036; 02044037; 02044038; 02044039; 02044041; 02044042; 02044043; 02044046; 02044050; 02044052; 02045001; 02045005; 02045012; 02045013; 02046013; 02046015; 02143003; 02143005; 02143007; 02143009; 02143010; 02144002; 02144005; 02144007; 02144009; 02144011; 02144013; 02144016; 02144018; 02144020; 02144022; 02144023; 02144025; 02144037; 02145001; 02145033; 02145039; 02145040; 02146027; 02146028; 02146029; 02146042; 02146048; 02147054; 02244054; 02244065; 02244071; 02245065; 02245067; 02245068; 02245069; 02245075; 02245081; 02245082; 02245087; 02245088; 02245089; 02245090; 02245104; 02245107; 02246051; 02246052; 02246055; 02246065; 02246127
6	00083437; 00083481; 00083592; 00083642; 01443004; 01444002; 01444003; 01445000; 01542006; 01544009; 01544011; 01544012; 01544013; 01544014; 01544016; 01544020; 01544021; 01546000; 01642003; 01642004; 01642010; 01642014; 01643001; 01643003; 01643004; 01643008; 01643015; 01643017; 01643018; 01644000; 01644001; 01644004; 01644005; 01644007; 01644010; 01644012; 01644013; 01644014; 01644015; 01644016; 01644018; 01644019; 01644023; 01645000; 01645003; 01645005; 01645007; 01645013; 01646000; 01646003; 01646004; 01647008; 01742001; 01742006; 01742014; 01742017; 01743002; 01744001; 01744002; 01744004; 01744006; 01744008; 01744010; 01744014; 01745000; 01745001; 01745007; 01746001; 01746006; 01746007; 01841001; 01841005; 01841015; 01842002; 01842004; 01842008; 01843005; 01844000; 01844003; 01844004; 01844009; 01844010; 01844015; 01845000; 01845021; 01941001; 01941004; 01941011; 01942001; 01942002; 01942006; 01942007; 01942008; 01942021; 01943042; 01943049; 01944018; 01944029; 01944059; 01945004; 01945039; 01947002; 01948003; 01949000; 02042000; 02042001; 02042008; 02042009; 02042010; 02042012; 02042015; 02042016; 02042017; 02042031; 02043005; 02043016; 02043024; 02043026; 02043027; 02044014; 02045020; 02046005; 02144014; 02144015; 02144024; 02145004; 02145012; 02145013; 02145014; 02145015; 02145016; 02145017; 02145018; 02145019; 02145020; 02145021; 02145022; 02145023; 02145024; 02145025; 02145026; 02145027; 02145028; 02145029; 02145030; 02145031; 02145032; 02145033; 02145034; 02145035; 02145036; 02145037; 02145038; 02145039; 02145040; 02145041; 02145042; 02145043; 02145044; 02145045; 02145046; 02145047; 02145048; 02145049; 02145050; 02145051; 02145052; 02145053; 02145054; 02145055; 02145056; 02145057; 02145058; 02145059; 02145060; 02145061; 02145062; 02145063; 02145064; 02145065; 02145066; 02145067; 02145068; 02145069; 02145070; 02145071; 02145072; 02145073; 02145074; 02145075; 02145076; 02145077; 02145078; 02145079; 02145080; 02145081; 02145082; 02145083; 02145084; 02145085; 02145086; 02145087; 02145088; 02145089; 02145090; 02145091; 02145092; 02145093; 02145094; 02145095; 02145096; 02145097; 02145098; 02145099; 02145100; 02145101; 02145102; 02145103; 02145104; 02145105; 02145106; 02145107; 02145108; 02145109; 02145110; 02145111; 02145112; 02145113; 02145114; 02145115; 02145116; 02145117; 02145118; 02145119; 02145120; 02145121; 02145122; 02145123; 02145124; 02145125; 02145126; 02145127; 02145128; 02145129; 02145130; 02145131; 02145132; 02145133; 02145134; 02145135; 02145136; 02145137; 02145138; 02145139; 02145140; 02145141; 02145142; 02145143; 02145144; 02145145; 02145146; 02145147; 02145148; 02145149; 02145150; 02145151; 02145152; 02145153; 02145154; 02145155; 02145156; 02145157; 02145158; 02145159; 02145160; 02145161; 02145162; 02145163; 02145164; 02145165; 02145166; 02145167; 02145168; 02145169; 02145170; 02145171; 02145172; 02145173; 02145174; 02145175; 02145176; 02145177; 02145178; 02145179; 02145180; 02145181; 02145182; 02145183; 02145184; 02145185; 02145186; 02145187; 02145188; 02145189; 02145190; 02145191; 02145192; 02145193; 02145194; 02145195; 02145196; 02145197; 02145198; 02145199; 02145200; 02145201; 02145202; 02145203; 02145204; 02145205; 02145206; 02145207; 02145208; 02145209; 02145210; 02145211; 02145212; 02145213; 02145214; 02145215; 02145216; 02145217; 02145218; 02145219; 02145220; 02145221; 02145222; 02145223; 02145224; 02145225; 02145226; 02145227; 02145228; 02145229; 02145230; 02145231; 02145232; 02145233; 02145234; 02145235; 02145236; 02145237; 02145238; 02145239; 02145240; 02145241; 02145242; 02145243; 02145244; 02145245; 02145246; 02145247; 02145248; 02145249; 02145250; 02145251; 02145252; 02145253; 02145254; 02145255; 02145256; 02145257; 02145258; 02145259; 02145260; 02145261; 02145262; 02145263; 02145264; 02145265; 02145266; 02145267; 02145268; 02145269; 02145270; 02145271; 02145272; 02145273; 02145274; 02145275; 02145276; 02145277; 02145278; 02145279; 02145280; 02145281; 02145282; 02145283; 02145284; 02145285; 02145286; 02145287; 02145288; 02145289; 02145290; 02145291; 02145292; 02145293; 02145294; 02145295; 02145296; 02145297; 02145298; 02145299; 02145300; 02145301; 02145302; 02145303; 02145304; 02145305; 02145306; 02145307; 02145308; 02145309; 02145310; 02145311; 02145312; 02145313; 02145314; 02145315; 02145316; 02145317; 02145318; 02145319; 02145320; 02145321; 02145322; 02145323; 02145324; 02145325; 02145326; 02145327; 02145328; 02145329; 02145330; 02145331; 02145332; 02145333; 02145334; 02145335; 02145336; 02145337; 02145338; 02145339; 02145340; 02145341; 02145342; 02145343; 02145344; 02145345; 02145346; 02145347; 02145348; 02145349; 02145350; 02145351; 02145352; 02145353; 02145354; 02145355; 02145356; 02145357; 02145358; 02145359; 02145360; 02145361; 02145362; 02145363; 02145364; 02145365; 02145366; 02145367; 02145368; 02145369; 02145370; 02145371; 02145372; 02145373; 02145374; 02145375; 02145376; 02145377; 02145378; 02145379; 02145380; 02145381; 02145382; 02145383; 02145384; 02145385; 02145386; 02145387; 02145388; 02145389; 02145390; 02145391; 02145392; 02145393; 02145394; 02145395; 02145396; 02145397; 02145398; 02145399; 02145400; 02145401; 02145402; 02145403; 02145404; 02145405; 02145406; 02145407; 02145408; 02145409; 02145410; 02145411; 02145412; 02145413; 02145414; 02145415; 02145416; 02145417; 02145418; 02145419; 02145420; 02145421; 02145422; 02145423; 02145424; 02145425; 02145426; 02145427; 02145428; 02145429; 02145430; 02145431; 02145432; 02145433; 02145434; 02145435; 02145436; 02145437; 02145438; 02145439; 02145440; 02145441; 02145442; 02145443; 02145444; 02145445; 02145446; 02145447; 02145448; 02145449; 02145450; 02145451; 02145452; 02145453; 02145454; 02145455; 02145456; 02145457; 02145458; 02145459; 02145460; 02145461; 02145462; 02145463; 02145464; 02145465; 02145466; 02145467; 02145468; 02145469; 02145470; 02145471; 02145472; 02145473; 02145474; 02145475; 02145476; 02145477; 02145478; 02145479; 02145480; 02145481; 02145482; 02145483; 02145484; 02145485; 02145486; 02145487; 02145488; 02145489; 02145490; 02145491; 02145492; 02145493; 02145494; 02145495; 02145496; 02145497; 02145498; 02145499; 02145500; 02145501; 02145502; 02145503; 02145504; 02145505; 02145506; 02145507; 02145508; 02145509; 02145510; 02145511; 02145512; 02145513; 02145514; 02145515; 02145516; 02145517; 02145518; 02145519; 02145520; 02145521; 02145522; 02145523; 02145524; 02145525; 02145526; 02145527; 02145528; 02145529; 02145530; 02145531; 02145532; 02145533; 02145534; 02145535; 02145536; 02145537; 02145538; 02145539; 02145540; 02145541; 02145542; 02145543; 02145544; 02145545; 02145546; 02145547; 02145548; 02145549; 02145550; 02145551; 02145552; 02145553; 02145554; 02145555; 02145556; 02145557; 02145558; 02145559; 02145560; 02145561; 02145562; 02145563; 02145564; 02145565; 02145566; 02145567; 02145568; 02145569; 02145570; 02145571; 02145572; 02145573; 02145574; 02145575; 02145576; 02145577; 02145578; 02145579; 02145580; 02145581; 02145582; 02145583; 02145584; 02145585; 02145586; 02145587; 02145588; 02145589; 02145590; 02145591; 02145592; 02145593; 02145594; 02145595; 02145596; 02145597; 02145598; 02145599; 02145600; 02145601; 02145602; 02145603; 02145604; 02145605; 02145606; 02145607; 02145608; 02145609; 02145610; 02145